EVALUACIÓN Y GESTIÓN DE LA CALIDAD EDUCATIVA

Un enfoque metodológico

Teresa González Ramírez (coord.)



EDICIONES ALJIBE

COLECCIÓN BIBLIOTECA DE EDUCACIÓN

Tlf.: 95 - 271 43 95
Fax: 95 - 271 43 42
Pavia, 8 - 29300-ARCHIDONA (Málaga)
e-mail: aljibe@indico.com

I.S.B.N.: 84-95212-83-8
Depósito legal: MA-1.102-2.000

Cubierta: Esther Morcillo y Fernando Cabrera.

Imprime: Imagraf. Málaga.

Teresa González Ramírez
(Coordinadora)

EVALUACIÓN Y GESTIÓN DE LA CALIDAD EDUCATIVA
Un enfoque metodológico

EDICIONES
ALJIBE



A nuestros alumnos

AUTORES

- ***Dra. Teresa González Ramírez (Coordinadora).***
- ***Dra. Leonor Buendía Eisman**. Catedrática del Área de Métodos de Investigación y Diagnóstico en Educación. Universidad de Granada.*
- ***Dra. Mª Pilar Colás Bravo**. Catedrática del Área de Métodos de Investigación y Diagnóstico en Educación. Universidad de Sevilla.*
- ***Dña. Elvira Congosto Luna**. Coordinadora del Departamento de Calidad, Evaluación Institucional y Formación del Profesorado del Centro Universitario Francisco de Vitoria.*
- ***Dr. José Manuel García Ramos**. Catedrático del Área de Métodos de Investigación y Diagnóstico en Educación. Universidad Complutense de Madrid.*
- ***Dra. Mercedes García García**. Profesora Titular del Área de Métodos de Investigación y Diagnóstico en Educación. Universidad Complutense de Madrid.*
- ***Dña. Beatriz García Lupión**. Profesora del Área de Métodos de Investigación y Diagnóstico en Educación. Universidad de Granada.*
- ***Dr. José Mª Gobantes Ollero**. Profesor del Área de Métodos de Investigación y Diagnóstico en Educación. Universidad de La Laguna.*
- ***Dra. Teresa González Ramírez**. Profesora del Área de Métodos de Investigación y Diagnóstico en Educación. Universidad de Sevilla.*
- ***Dr. Jesús Miguel Muñoz Cantero.** Profesor Titular del Área de Métodos de Investigación y Diagnóstico en Educación. Universidad de La Coruña.*

ÍNDICE

PRESENTACIÓN

"Lo que no se define no se puede medir.
Lo que no se mide no se puede mejorar.
Lo que no se mejora se degrada siempre"
(Lord Kelvin, 1885)

Vivimos en una sociedad compleja, interdependiente y cambiante, caracterizada por la rapidez en la creación y difusión del conocimiento. Esta necesidad de la sociedad actual de acomodarse a los cambios hace que ponga sus puntos de mira en los resultados de la educación, de tal manera que actualmente hemos pasado de hablar del "*gasto formativo*" a la "*inversión formativa*" como uno de los retos fundamentales que se plantean a las sociedades industrializadas al inicio del siglo XXI. Esta ha sido una de las principales conclusiones del 56º Congreso del Instituto Nacional de Hacienda Pública (IIPF) celebrado en nuestra Universidad con el objetivo fundamental de conocer el impacto que los recursos públicos ejercen en la mejora del capital humano.

De otra parte, la historia de la evaluación nos muestra -en épocas pasadas porque las políticas económicas han centrado sus esfuerzos en la acumulación del capital tangible y en la actualidad porque todo parece apuntar que los esfuerzos van a dirigirse hacía la formación de capital humano- que cuando empezamos a cuestionar o replantear qué aportan los sistemas educativos a la sociedad en general, la evaluación adquiere un gran valor estratégico; se convierte en la herramienta que nos permite gestionar la calidad de la educación. Como señala House (1992) la evaluación estructuralmente se ha ido integrando mas en el funcionamiento organizativo de las instituciones y conceptualmente se ha flexibilizado.

Estas circunstancias, externas al propio campo disciplinar, han hecho que la evaluación en nuestro país, esté empezando a desempeñar un papel de primer orden; así tras la amplia experiencia adquirida durante décadas en lo relativo a la evaluación de los aprendizajes alcanzados por los alumnos y a la medición de las diferencias existentes entre ellos, el siguiente ámbito abordado de manera rigurosa fue el de la evaluación de programas educativos. Pero el cambio fundamental de orientación científica y práctica se produjo cuando las instituciones educativas y el propio sistema en su conjunto comenzaron a ser objeto de evaluación, de tal manera que en la actualidad no podemos desprendernos de un concepto de evaluación emergente, pluralista, en estrecha vinculación con una multiplicidad de contextos en los que se han generado prácticas evaluativas también muy diversas.

En la política educativa destacan el énfasis puesto en la evaluación del aprendizaje de los alumnos, de programas y sistemas educativos así como la generalización de la evaluación de centros educativos e instituciones y de los profesionales de la educación (evaluación de la formación). Este crecimiento exponencial en la cultura contemporánea de la evaluación, aplicada a una multiplicidad de situaciones evaluativas justifica por una parte, la necesidad de una profesionalización de los evaluadores y de otra, el reconocimiento de la evaluación como disciplina con entidad propia, dotada de metodología y teoría propias.

Esta publicación dirigida a profesores y alumnos de las Facultades de Ciencias de la Educación y a todos aquellos profesionales que se dediquen al campo de la evaluación en contextos formativos tiene el propósito fundamental de presentar tanto el campo científico-metodológico de la evaluación así como los diferentes contextos evaluativos que han ido incorporándose a esta disciplina. Bajo este propósito, pretende tres objetivos fundamentales:

a) Recapitular las principales aportaciones sobre evaluación y calidad desde los distintos ámbitos de evaluación y gestión de la calidad educativa en los que se está trabajando.
b) Servir de guía y orientación a profesores, alumnos y profesionales de la evaluación tanto en el plano teórico como metodológico. Los diversos ámbitos evaluativos expuestos tienen una dimensión didáctica y aplicativa.
c) Proporcionar una herramienta para el diseño y desarrollo de estudios evaluativos aplicados a diferentes ámbitos de gestión de la calidad educativa (evaluación de programas, de sistemas educativos, de la universidad, de centros, etc.) El contenido teórico y aplicativo de la publicación aporta referentes importantes para su posterior aplicación en contextos académicos y profesionales diversos.

Puesto que los destinatarios de este libro no sólo necesitan conocer los aspectos teóricos de la evaluación sino también su vertiente aplicativa, el texto se estructura en dos partes: la primera, "*Evaluación y Calidad Educativa*" presenta las principales coordenadas que han regido la evolución de la evaluación desde un

punto de vista científico y metodológico y en estrecha conexión con el concepto de calidad educativa; la segunda está dirigida a presentar "las formas de hacer de la evaluación" en los distintos ámbitos evaluativos tratados: evaluación de programas, de profesores, de centros, evaluación institucional universitaria y evaluación de sistemas educativos. La secuencia en la que se presentan no es aleatoria; cuando empezamos a hablar de evaluación el único ámbito evaluativo en desarrollo era la evaluación de programas, de tal manera que en la bibliografía de la época evaluación, evaluación de programas e investigación evaluativa se trataban de manera intercambiable, han sido las prácticas evaluativas generadas en diferentes contextos las que han ido diversificando el campo conceptual de la evaluación. En la actualidad, cada uno de estos ámbitos no puede separarse del concepto de calidad educativa. Se presenta en siete capítulos que quedan estructurados como sigue:

En el *Capítulo I* se expone una concepción global y general del panorama científico-metodológico de la evaluación educativa actual. Desde esta perspectiva, se realiza un breve recorrido por las principales áreas de la medición educativa. La evaluación desde un punto de vista científico, va estrechamente ligada a la medición. A este recorrido histórico la autora añade las tendencias científicas actuales así como el impacto de las tecnologías en las prácticas evaluativas. Termina este capítulo argumentando nuevos retos y perspectivas que tiene planteada la evaluación educativa: cambios que afectan a los aprendizajes, uso de todas las formas educativas que ofrecen la tecnología de las comunicaciones y sistemas multimedia y diseños de evaluación en muy distintos formatos y dirigidos a muy distintas funciones.

El *Capítulo II* en estrecha conexión con el anterior, se centra en aquellas dimensiones externas que desde fuera también han ido moldeando el campo de la evaluación. El énfasis que en los últimos años se ha puesto en el concepto de calidad educativa no es mas que el resultado de cómo han ido evolucionando las estructuras educativas y/o formativas. Esta evolución motivada por el control que ejercen en las políticas públicas los distintos agentes que en ellas intervienen -administración, estamento académico y usuarios- ha generado la necesidad de "rendir cuentas" al sistema educativo. En esta nueva era de la evaluación el movimiento de "accountability" adquiere nuevos matices por la necesidad no sólo de conocer los resultados de la educación sino de aprender a gestionar los procesos. Desde este nuevo enfoque los modelos de gestión de la calidad empiezan a ser considerados en el ámbito educativo y se ponen en marcha medidas gubernamentales que desarrollan políticas de calidad que tienen una incidencia clara en la manera de gestionar los procesos y los resultados en las instituciones educativas y de los sistemas educativos en general. En definitiva, es el marco social en el que nos movemos el que está pidiendo que optimicemos los procesos y mejoremos los resultados de la educación.

El *Capítulo III*, es el que sirve de inicio y presentación a los distintos ámbitos de gestión de la calidad educativa tratados. El autor plantea la *Evaluación de Programas* desde la idea de que los programas educativos son elaboraciones sociales

que responden a necesidades, intereses y finalidades diversas, determinadas socialmente. Desde esta perspectiva el programa a evaluar tendrá la impronta del contexto -político, económico, administrativo o social- en el que se desarrolla y de las personas involucradas en él de una u otra forma. Esta concepción de partida, aporta nuevas ideas a nivel metodológico: son los contextos en los que se desarrolla la evaluación de programas los que nos invitan a la creación metodología fuera de todo apriorismo que mediatice la naturaleza del objeto a evaluar.

El *Capítulo IV* plantea la *Evaluación del Profesorado* como proceso sistemático e integrado en la actividad educativa, que ayuda a comprender la realidad, invitando a la reflexión y como facilitadora de las decisiones adecuadas para el desarrollo profesional y, consecuentemente, para la mejora institucional. Es el ámbito evaluativo que desde sus inicios ha tenido un referente legislativo. En el ámbito universitario, la Ley de reforma Universitaria (LRU, 1983) contempla la evaluación del profesor, pero se centra, a diferencia de la LOGSE, en la individualidad del profesor como objeto de evaluación y no dentro del contexto de una evaluación institucional. Desde esta situación de partida, el capítulo aborda diferentes modelos de evaluación del profesorado, plantea diferentes agentes que intervienen en la evaluación del mismo y estudia los requisitos que deben contemplar las técnicas e instrumentos para la evaluación del profesor. Finalmente analiza las características diferenciales en los procesos de evaluación en función del nivel o etapa educativa de que se trate, con especial detenimiento en el caso universitario.

El *Capítulo V* sobre *Evaluación y Gestión de la Calidad de los Centros Educativos* se desarrolla sobre un concepto de centro educativo entendido como "unidad organizativa funcional" en el que el concepto de evaluación se aplica al centro en su totalidad. Desde esta perspectiva se presentan distintos modelos de evaluación de centros entendiendo éste como paso previo para la realización de un estudio evaluativo y su implicación en el desarrollo de una teoría. El autor presenta los modelos de evaluación en primer lugar desde el punto de vista de las organizaciones, después atendiendo a un criterio de orden epistemológico, y finalmente partiendo del criterio de calidad. Termina este capítulo presentando la secuencia metodológica para desarrollar un plan de evaluación de centros.

El *Capítulo VI* destinado a la *Evaluación Institucional Universitaria* se presenta como uno de los ámbitos evaluativos que en los últimos años se ha convertido en tema prioritario de la política educativa en toda Europa. Las exigencias externas ejercidas sobre la institución universitaria adoptando para ellos distintos discursos y modelos formativos sobre cuáles son sus principales metas y objetivos, han llevado a la mayoría de los países desarrollados a poner en marcha sistemas de evaluación institucional. En este capítulo se expone el recorrido de la evaluación institucional en España a partir del Plan Nacional de Evaluación de la Calidad de las Universidades. Finalmente recoge una síntesis comparativa de los modelos de evaluación institucional en diferentes países.

La *Evaluación de Sistemas Educativos*, cierra en el *capítulo VII* los distintos ámbitos evaluativos tratados. Este capítulo plantea los antecedentes de la evaluación de sistemas educativos, la relevancia y usos de la evaluación, así como distintos enfoques para la evaluación de sistemas educativos -vinculado a la estructura de gestión y funcionamiento, o vinculado al enfoque de calidad educativa-. Dentro de este capítulo tiene un tratamiento especial, por su relevancia en la actualidad, el tema complejo de los indicadores de la calidad de la educación. Esta ha sido una de las líneas prioritarias de actuación del INCE desde su creación, la reciente publicación sobre los indicadores de la educación (INCE, 2000) es una muestra de la sensibilidad que existe por la creación de un sistema de indicadores que nos ayude a interpretar la información obtenida en redes de conocimiento con un valor explicativo suficiente sobre el funcionamiento del sistema educativo español. Finaliza este capítulo desarrollando la evaluación del sistema educativo en España. La creación del Instituto Nacional de Calidad y Evaluación (INCE), ha supuesto una plataforma dentro de las políticas de calidad del Ministerio de Educación y Cultura para ejercer un control y seguimiento sobre el mismo.

En síntesis, los contenidos presentados a lo largo de estos siete capítulos suponen un recorrido en el tiempo sobre el campo disciplinar de la evaluación. Esperamos que el lector encuentre en esta publicación una guía útil para aproximarse a los contextos de la evaluación como fuente de conocimiento y de desarrollo práctico, y poder contribuir así de esta manera a la mejora de la calidad de la educación entendida como esfuerzo sistemático y sostenido para cambiar las condiciones internas y externas que rigen los contextos educativos.

Finalmente, no quisiera terminar sin expresar mi agradecimiento a todos los autores que han participado en esta publicación. La autoría conjunta siempre comporta riesgos, pero la colaboración y entrega que han reinado durante su desarrollo ha sido una muestra del espíritu compartido para poder abrir así un espacio de reflexión y comunicación entre los profesionales de la evaluación.

Teresa González Ramírez
Sevilla, Julio de 2000

PRIMERA PARTE

EVALUACIÓN Y CALIDAD EDUCATIVA

CAPÍTULO I

EVALUACIÓN EDUCATIVA: PANORAMA CIENTÍFICO Y NUEVOS RETOS

Mª Pilar Colás Bravo.[1]
Universidad de Sevilla

OBJETIVOS

1. Construir una concepción global y general del panorama actual de la evaluación educativa
2. Realizar una aproximación conceptual y comprensiva a algunas líneas de investigación en el área de la medición
3. Identificar las principales áreas científicas de investigación en evaluación
4. Conocer aplicaciones de los test a la educación y su problemática
5. Conocer las tendencias científicas actuales en relación a los test
6. Tomar conciencia del impacto tecnológico en los procedimientos de evaluación
7. Obtener una visión comprensiva de las perspectivas y nuevos retos que se plantean en la evaluación educativa
8. Propiciar una comprensión de la evaluación en su dimensión científica

CONTENIDOS

---

[1] Facultad de Ciencias de la Educación. Área de Métodos de Investigación y Diagnóstico en Educación. Av. San Francisco Javier s/n. 41005-SEVILLA

INTRODUCCIÓN

El momento actual de la evaluación educativa se caracteriza y explica por la confluencia fundamentalmente de intereses y desarrollos científicos y sociales. Estas dos dimensiones interaccionan nutriéndose y potenciándose mutuamente como expresamos en el gráfico 1.

Hoy día la evaluación de la calidad educativa posee un gran interés político, social y económico. Por ello la evaluación ha de entenderse e interpretarse bajo estos referentes. Por otra parte la evolución de las teorías científicas y de sus propios marcos conceptuales hacen posible determinados tipos de prácticas evaluativas. También el impacto tecnológico propicia una forma de hacer distinta las actividades propias de la evaluación. Por tanto, los modelos teórico-científico o los desarrollos técnicos y las tecnologías de la información y la comunicación, aportan y proveen de herramientas conceptuales y técnicas para cumplir estos cometidos.

El siguiente gráfico expresa las relaciones entre las aportaciones científicas y las demandas del contexto social.

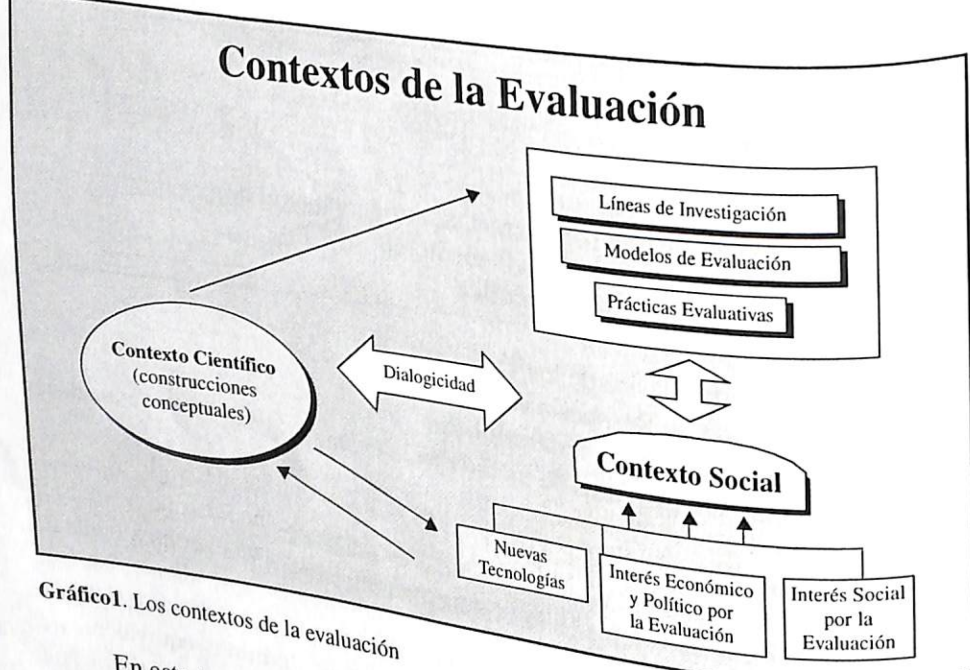


Gráfico1. Los contextos de la evaluación

En este texto se abordan las dos perspectivas: la social-contextual y la científica. Este capítulo se dedica exclusivamente a hacer un breve recorrido por la dimensión más científica y/o técnica de la evaluación. Los restantes capítulos muestran diferentes contextos en los que tienen lugar las prácticas evaluativas.

1. LA EVALUACIÓN EN EL CONTEXTO EDUCATIVO

La evaluación en el contexto educativo sigue un proceso evolutivo a lo largo de la historia que de forma genérica expresamos en el siguiente gráfico:

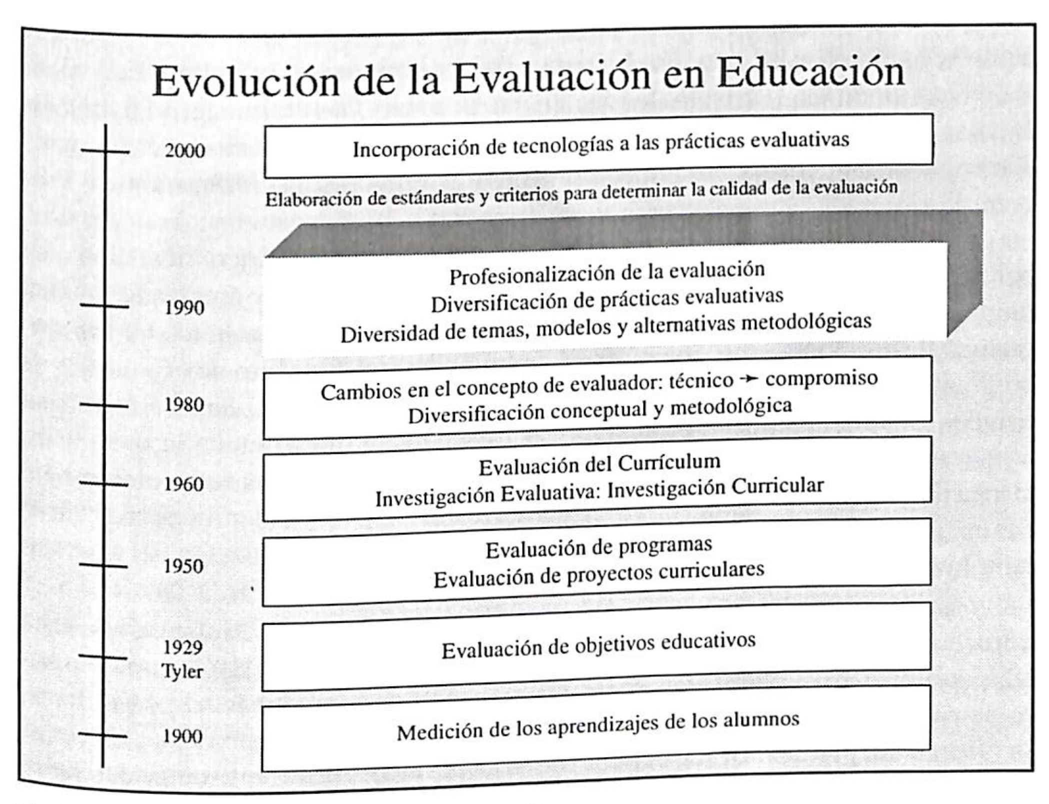


Gráfico 2. Evolución de la evaluación en Educación.

Este cuadro sintetiza el desarrollo histórico y los cambios que se producen en los distintos periodos históricos.

En un primer estadio, que podría situarse a comienzos del siglo veinte, la evaluación se circunscribe al desarrollo de instrumentos de medida que posibiliten la evaluación del alumno. Las primeras prácticas evaluativas en el contexto educativo se vinculan por tanto a la valoración de resultados de aprendizaje. En este periodo la evaluación va estrechamente ligada al concepto de *medición* del rendimiento. Evaluación y medición constituyen, por tanto, términos equivalentes e intercambiables. En este periodo la medición (evaluación) no se vincula ni con programas educativos ni con el currículum escolar, sino que constituye una proyección científica de las prácticas científicas de la época, asentadas en el paradigma positivista.

En un segundo momento histórico, situado en el segundo cuarto del siglo veinte, la evaluación se abre a otras dimensiones educativas. Tyler, en 1929, comienza a considerar la evaluación con referencia a objetivos planteados en el diseño curricular. Aunque Tyler evaluó fundamentalmente el aprendizaje, realizó una importante aportación señalando que el término evaluación implicaba algo más que la mera aplicación del test, y habría de servir para determinar hasta qué punto se habían conseguido los objetivos establecidos en el currículum. Este planteamiento modifica, por tanto, la idea inicial de medir las diferencias individuales en cuanto a capacidades y cualidades de los alumnos y la sitúa en términos de diseño curricular. Este enfoque se considera al inicio de la Evaluación de Programas.

Hacia los años 50 resurge la idea de la evaluación de programas, motivada por la necesidad de evaluar proyectos curriculares de distintas áreas que reciben fuertes subvenciones estatales en Estados Unidos. En la década de los 60 los programas alcanzan una proyección social, ampliándose el concepto de evaluación de programas más allá de la faceta estrictamente académica y/o escolar, originando en Estados Unidos una nueva profesión. En Europa, es Gran Bretaña la que en los años sesenta financia proyectos para la evaluación de programas curriculares, bajo la idea de que la evaluación debía servir para ayudar a quienes planificaban el currículum. Este tipo de evaluación adopta la denominación de Evaluación del Currículum, Investigación Curricular o Investigación Evaluativa.

La década de los 80 supone una diversificación, tanto desde el punto de vista estructural, como conceptual. Conceptualmente se produce un desplazamiento hacia posiciones más pluralistas, tanto en los métodos o prácticas que se utilizan, como en los modelos que las guían. Uno de los cambios más específicos y claves es su orientación política. Desde el punto de vista político, el concepto de evaluador se ha desplazado desde el de "técnico" o "experto", al de profesional que desempeña una actividad con cierto impacto político. La evaluación pasa de considerarse una actividad técnica cuyo cometido es "*rendir cuentas*", a adoptar un compromiso o posición política. En la década de los noventa se consolida el interés político por actividades evaluativas en muy distintos ámbitos y contextos: educativos, sociales, políticos, institucionales, etc. La evaluación por tanto se reconoce y advierte de gran interés a nivel científico, político y social.

El momento actual se caracteriza por las siguientes notas (Arias, Verdugo y Rubio, 1995: 28):

- ***Profesionalización*** de la evaluación que se traduce en la existencia de centros y actividades formativas destinadas a formar profesionales especializados (Universidades, estudios de postgrado, masters, etc.), y la creación de asociaciones y estándares profesionales, como la APA (*American Psychological Association*), que elabora stándares para su uso en estudios, o el Joint Commitee Evaluation Standards, que desarrolla periódicamente criterios para su aplicación en evaluaciones, todo junto con la proliferación de publicaciones especializadas.
- ***Diversificación de prácticas evaluativas*** que quedan reflejadas en la diversidad de profesionales que asumen estas funciones en cuanto a sus campos de procedencia (educación, economía, salud, sociología, psicología), y la diversidad de programas y metodologías que se aplican.

La diversidad estructural se manifiesta en la ampliación de la *evaluación a centros, instituciones y sistemas educativos,* hecho que puede ser entendido por el interés de las políticas educativas, tanto a nivel local como nacional e internacional, ya que se considera la evaluación una herramienta importante para conocer los resultados educativos y plantear reformas educativas (Glaser y Silver, 1994).

Una de las parcelas que últimamente está recibiendo gran atención es la *evaluación de sistemas educativos.* La identificación y medición de indicadores de calidad de sistemas educativos, centros e instituciones constituye un reto metodológico en esta línea. Paralelamente a la extensividad de la territorialidad de la evaluación crecen cuestionamientos y reflexiones acerca de la naturaleza y formas de evaluación, así como sobre los usos que se hacen de ella. Algunos de estos planteamientos pueden verse en Moss (1996) y Schmidt, Jakwerth y McKnight (1998) o Wolf (1998).

2. LA EVALUACIÓN DESDE UNA PERSPECTIVA CIENTÍFICA Y METODOLÓGICA

Las aplicaciones evaluativas corren paralelas al trabajo científico generado en el campo científico de la medición. La *evaluación*, desde un punto de vista científico, va estrechamente ligada a la *medición.*

2.1. Aproximación conceptual a la medición

La medida es un componente esencial en cualquier práctica evaluativa, ya que permite fundamentar los juicios o valoraciones de aquello que se evalúa sobre datos empíricos rigurosos, válidos y fiables. En la medición, por tanto, ocupa un lugar tradicional y privilegiado la construcción de instrumentos, *Tests*, que tienen por objeto la asignación de números a atributos o conductas de las personas. Otra faceta en la que la medición también se proyecta es en la construcción de *escalas* para la medición de estímulos originando el escalamiento. Estas áreas de trabajo comparten un interés central, la "traducción" de aspectos observables considerados relevantes para la evaluación referidos a personas, contextos, situaciones, etc. a un sistema formal. Los datos frecuentemente se registran en lenguaje matemático.

Lord y Novick (1968: 17), definen la medida como "procedimientos para la asignación de números (puntuaciones, medidas) a propiedades específicas de uni-

dades experimentales, de modo que caractericen y preserven las relaciones especificadas en el dominio comportamental". De forma equivalente Stevens (1951: 22) entiende que medida es la "asignación de números a objetos o sucesos según ciertas reglas. Estas reglas suponen preservar un isomorfismo entre las características del sistema numérico y las propiedades de los objetos que se miden".

El desarrollo de reglas sistemáticas y unidades significativas de medida para cuantificar las observaciones empíricas se denomina *escalamiento*. Una escala de medida se establece cuando se define el conjunto de valores posibles que pueden asignarse, y se establece la regla de asignación que desarrolla la correspondencia entre el sistema empírico y el sistema numérico. El escalamiento, a partir de la publicación de los trabajos de Stevens (1946, 1951), se abre como área de interés para los científicos sociales. Actualmente es la Teoría Representacional de la Medición la posición más ortodoxa en cuanto a la conceptualización de la medida. Esta teoría axiomática y formalizada aborda la medida en torno a tres grandes áreas: la representación, la unicidad y la significación.

Desde el punto de vista de *la representación* uno de los conceptos centrales es el sistema relacional. Y desde este punto de vista "la medición supone encontrar un sistema relacional numérico con una estructura semejante al sistema relacional empírico que se pretende medir (...). Por tanto, medir supone encontrar un sistema relacional numérico que pueda representar al sistema relacional empírico de interés" (Martínez Arias, 1995: 24). Las distintas opciones y posibilidades dan origen a diferentes escalas de medida. Así, se pueden representar relaciones de identidad o de diferencia a través de reglas de asignación que dan lugar a escalas de *medida nominal*. Los números también pueden representar relaciones de orden, dando lugar las reglas de asignación a una escala de *medida ordinal*. La representación de la igualdad de diferencias entre pares de objetos con relación a alguna propiedad o atributo da lugar a una *escala de intervalo*, y por último el cuarto nivel de medida, la *medida de razón*, asume un cero fijo, lo que en el sistema empírico supone un cero absoluto.

El problema de la *unicidad* hace referencia a la arbitrariedad de los números elegidos según la Teoría Representacional. Las relaciones numéricas que se pueden emplear para la representación de un sistema empírico pueden ser diferentes manteniendo el isomorfismo, es decir, pueden obtenerse distintas escalas de números para la misma variable o atributo. La solución al problema de la unicidad se halla en la posibilidad de intercambiar conjuntos de números concretos por otros, sin que se modifiquen las relaciones numéricas que representan a las relaciones empíricas.

La significación hace referencia a la validez de una conclusión empírica inferida a partir de una conclusión numérica. Esta validez siempre es relativa al tipo de escala en el que se basan las inferencias. Stevens (1946), planteó la solución al problema en términos de estadísticos admisibles para cada tipo de escala, basándo-

se en la invarianza o transformaciones admisibles de la escala. Así, por ejemplo, la media y otros estadísticos basados en ella como la varianza o el coeficiente de correlación lineal no deberían utilizarse con escalas ordinales, ya que los cálculos demandan más relaciones que las de orden.

La medición, por tanto, se ocupa de establecer y valorar los procesos de transformación de las observaciones o registros obtenidos de la realidad en categorías numéricas o cualitativas. El eje central en el que se proyecta la medición se refiere a la conexión de constructos científicos con fenómenos observables y empíricos.

La medición en su proyección educativa incluye recoger información cuantitativa sobre alumnos, profesores o escuelas, por lo que constituye evidentemente un tema muy amplio. En él tiene cabida tanto la creación y la evaluación de sistemas de observación en clase como el diseño y uso de cuestionarios (Linn, 1988: 148). En esta misma línea, se expresan los catedráticos Hernández Pina y Del Rincón.

Hernández Pina (1995: 138), plantea que la medición ha de entenderse de forma amplia, identificándose con el acto de recogida de información. Para Del Rincón y otros (1995: 48), "la medición sirve para establecer y valorar críticamente las estrategias a seguir en el proceso de transformar las categorías observadas en los fenómenos en categorías numéricas o cualitativas".

2.2. La medición como área científica

La primera etapa científica de la medición se identifica como "testing". Ésta se caracteriza por el interés en recabar medidas de características o dimensiones individuales, para establecer diferencias entre los sujetos de un colectivo (normativo). Por tanto el objeto de la medición son las características individuales de los sujetos. Guba y Lincoln (1989), denominan a esta etapa "evaluación de primera generación". La sitúan a comienzos de siglo y la describen como la generación de la medida. Su contexto epistemológico se caracteriza por el florecimiento de las corrientes posivistas y empiristas que marcan el interés por obtener medidas objetivas, a lo que se une el desarrollo de métodos estadísticos y el proceso de industrialización que demanda selección y acreditación de fuerza de trabajo cada vez más especializada.

La psicometría surge como conjunto de métodos y técnicas implicadas en la medición de variables psicológicas. Trata todo lo relacionado con la medición, poniendo énfasis y especializándose en las propiedades métricas exigibles a las mediciones, independientemente del campo sustantivo de aplicación, y de los instrumentos utilizados. La validez y fiabilidad de las mediciones constituyen, por tanto, temáticas relevantes de la psicometría. Este campo disciplinar se articula en torno a cinco grandes bloques temáticos: Teoría de la Medición, que abarca todo lo relativo a la fundamentación teórica de la medida; Teoría de los Tests, que aborda la lógica

y los modelos matemáticos subyacentes a la construcción y uso de los tests: Escalamiento Psicológico, que aborda la problemática inherente al escalamiento de estímulos, y Técnicas Multivariadas, que junto con la Tecnología Estadística, se aplican en la construcción y análisis de los instrumentos de medida. Cada una de estas áreas temáticas están, a su vez, altamente especializadas y estructuradas en subáreas. Fruto del desarrollo disciplinar de este último campo, a partir de los años sesenta, se populariza el término de Psicología Matemática que estudia y se especializa en modelos matemáticos elaborados para áreas específicas de la psicología. Ambas disciplinas encuentran frecuentes solapamientos en sus contenidos (Muñiz, 1994).

En el siguiente gráfico se recogen las parcelas científicas que se desarrollan a lo largo del devenir científico de la medición.

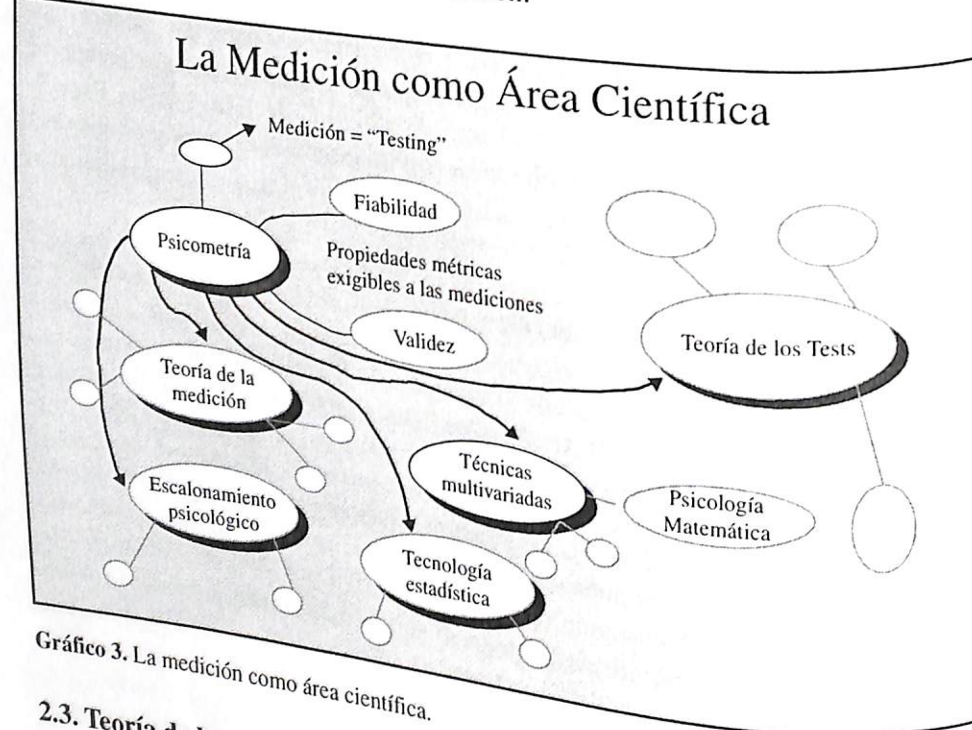


Gráfico 3. La medición como área científica.

2.3. Teoría de los Test

Uno de los temas centrales y tradicionales de la medición es la *construcción de test*. Sobre ellos se ha generado la Teoría de los Test que supone trasladar la problemática de la medida a la clasificación, situación o diferenciación de los sujetos en rasgos generalmente no observables directamente, sino a través de la proyección empírica de constructos. La Teoría de los Test se interesa por el desarrollo de modelos matemáticos, útiles para el análisis de datos proporcionados por los suje-

tos. También aborda y se plantea como campo de estudio toda la problemática que se genera en torno a la elaboración de instrumentos de medida en el campo de la educación y la psicología, tratando específicamente problemáticas tales como las que indica Martínez Arias (1995: 24):

a) *No existe una única aproximación a la medida de un constructo* que sea universalmente aceptada. Existe siempre la posibilidad de que dos teóricos seleccionen diferentes tipos de conducta para la definición operativa de un mismo constructo.
b) *Las medidas educativas y psicológicas están basadas en muestras limitadas de conductas*. Delimitar el número de elementos y la variedad de expresiones en que se manifiesta un constructo, es uno de los principales problemas en el desarrollo de un instrumento de medida.
c) *La medida obtenida siempre tiene error.* Al estar basada en muestras limitadas en contenidos y recogidas en un momento en el tiempo, está afectada por todos los errores del muestreo.
d) *Falta de escalas con origen y unidades de medida bien definidas.*
e) *Dificultad de identificar un constructo con total independencia así como sus definiciones operativas.*

La Teoría de los Test aborda los modelos lógicos y matemáticos que subyacen a la construcción y uso de los test, siguiendo el modelo estadístico en que se basan para explicar el significado de las puntuaciones y hacer inferencias a partir de las puntuaciones obtenidas. En la Teoría de los Tests se identifican tres orientaciones básicas:

- Teoría Clásica de los Test.
- Teoría de la Generalizabilidad.
- Teoría de Respuesta al Ítem.

2.3.1. *Teoría Clásica de los Test*

La Teoría Clásica de los Test ha sido el modelo dominante en la Teoría de los Tests durante gran parte del siglo y, aún hoy día, tiene una vigencia más que notable en el campo de la práctica de la evaluación psicológica y educativa. Esta teoría arranca de los primeros trabajos de Spearman (primera década del siglo XX), y se desarrolla en conjunción con la Teorías de la Inteligencia. Spearman es el primero en formular una teoría psicológica de la inteligencia. En esta Teoría se plantea un modelo estadístico que fundamenta las puntuaciones de los tests, y permite la estimación de los errores de medida asociados a todo proceso de medición. El modelo lineal clásico que asume, plantea que la puntuación empírica de un sujeto en un test (x), consta de dos componentes aditivos; uno, la verdadera "puntuación" del sujeto en el test (v), y otro, el error (e) que inevitablemente va asociado a todo proceso de medición. A partir de este modelo, y unas asunciones mínimas, la teoría clásica

desarrolla todo un conjunto de deducciones encaminadas a estimar la cuantía del error que afecta a las puntuaciones de los tests.

Los años siguientes suponen una actividad psicométrica importante, tanto en el campo de la teoría como en el de la construcción y tecnología de los tests. En 1936 Guilford publica *Psychometric Methods* manual que sintetiza lo fundamental de los campos de trabajo de aquella época; Teoría de los Tests, Escalamiento Psicológico y Psicofísico. En esta época se funda la Sociedad Psicométrica Americana y se crea la revista *Psychometrika*, sucediéndose pocos años después otras publicaciones periódicas como *Educational and Psychological Measurement* (1941), *Journal of Educational Measurement* (1964) y *Applied Psychological Measurement* (1977). En los años 40 Stevens publica su famoso trabajo sobre escalas de medida, que obliga a plantearse el estatus teórico de las mediciones en las Teorías de los Tests, además de sus propiedades empíricas.

El escalamiento es el contenido del texto de Torgerson (1958), *Theory and Methods of Scaling*. En esta época los campos de la medida se enriquecen y diversifican. En 1968 aparece el libro de Lord y Novick, *Statistical Theories of Mental Test scores*, que sintetiza y analiza críticamente todo lo hecho anteriormente en la Teoría Clásica de los Test, abriendo nuevas y prometedoras perspectivas. En este libro se incluyen los Modelos de Rasgo Latente, que abrirán una línea completamente nueva, conocida hoy como Teoría de Respuesta a los Ítems (TRI). Este texto, por tanto, marca un antes y un después en la Teoría de los Tests.

2.3.2. *Teoría de Respuesta al Ítem*

Con el objeto de sintetizar los avances acumulados sobre distintos aspectos de los nuevos modelos, Lord (1980) estudia y recoge la producción científica de una década. Su libro abre la década de los ochenta con la especificación de la TRI y la revitalización de áreas como: bancos de ítems, sesgos, o test referidos al criterio, por citar algunos. Este planteamiento se nutre también del progreso tecnológico, los avances teóricos, y la aplicación profesional.

Como aportación diferenciada de la Teoría Clásica de los Tests, la TRI (Teoría de Respuesta al Item) pretende subsanar las limitaciones de que los resultados obtenidos de la aplicación de un test dependen de las circunstancias de aplicación, de las características del test y de los sujetos examinados (Álvaro, 1993 y García Cueto, 1993).

La contribución de la TRI, desde el punto de vista teórico de la medición, se centra en la posibilidad de obtener medidas invariantes respecto de los instrumentos utilizados y de los sujetos. Aporta también marcos conceptuales como es la función de información, bancos de ítems, etc, que suponen avances importantes. Según Muñiz y Hambleton (1992: 52), la función de información es sin duda una de las aportaciones más relevantes y supone un cambio radical para la construcción

y análisis de los instrumentos de medida, ya que permite conocer para qué valores es más preciso el test (o el ítem), con lo que el uso de un test u otro va a depender del tipo de sujetos a evaluar. No se utiliza el mismo test con sujetos de competencia baja, media o alta, lo cuál sólo es posible, gracias a la invarianza de las mediciones respecto del test utilizado que permite la TRI. Navas (1994) subraya la capacidad de los modelos TRI para dar soluciones potentes a una gran variedad de problemas psicométricos aplicados. En esta aportación ilustra, de forma breve, algunas de las aplicaciones más interesantes de la TRI, poniendo especial énfasis en aquellas en las que esta teoría realiza contribuciones novedosas en relación a los planteamientos del modelo clásico.

El momento actual se identifica con una hegemonía clara de los modelos TRI, como puede observarse en las aportaciones a congresos, en las publicaciones periódicas o en los monográficos. Distintas áreas se han visto impulsadas por la TRI, surgiendo en la producción científica textos monográficos sobre aspectos específicos de los mismos. Una más detallada historia de la TRI puede leerse en Muñiz y Hambleton (1992).

2.3.3. *Teoría de la Generalizabilidad*

Podemos decir por tanto, con relación a la Teoría de los Test que el panorama actual de este siglo se dibuja en torno a estas dos grandes teorías referidas y la Teoría de la Generalizabilidad. Esta teoría se considera una extensión de la TCT, e intenta resolver algunos de los problemas que se plantean desde la Teoría Clásica, mediante la aplicación de procedimientos derivadas del Análisis de Varianza y los Diseños Experimentales, poniendo el acento en la estimación de los diversos componentes de la varianza. Esta teoría se debe a Cronbach y colaboradores (1972). Posteriormente otros autores consolidan sus aportaciones. La teoría G avanza la idea de que el error de medida puede ser analizado, al menos en parte, distinguiendo entre componentes diversos, cada uno de ellos asociado con algún aspecto manipulable de una prueba: su contenido, las condiciones de aplicación, o juez. Es decir, esta teoría permite tener en cuenta, de modo explícito, las múltiples fuentes de error de medida. Según Blanco (1993: 186), la aportación de la Teoría G es ofrecer un marco más satisfactorio para estimar la fiabilidad y los márgenes de error. La Teoría G desarrolla sus propios conceptos tales como universo de generalización, faceta, coeficiente de generalizabilidad, etc. La Teoría G se aplica en Ciencias de la Educación a trabajos sobre la fiabilidad de los comportamientos de maestros en clase, la evaluación de la enseñanza por los alumnos, la evaluación de prerrequisitos (matemáticas, lengua,...) para el ejercicio de ciertas profesiones, etc. Aplicaciones de la Teoría G en el ámbito español las encontramos en Buendía y Pegalajar (1990) y Blanco (1993).

3. LOS TESTS COMO INSTRUMENTOS DE MEDIDA EN LA EVALUACIÓN

Los Tests constituyen un núcleo importante en esta primera etapa de la medida, a la vez que es el *leiv motiv* de áreas y desarrollos de la Psicometría, que ha aportado un gran bagaje de conocimientos con relación a la Teoría de la Medición, Teoría de los Tests, etc. El desarrollo de estos campos corre parejo y está marcado, a su vez, por la evolución de tests concretos.

En este período medición y evaluación se plantean como términos equivalentes, en tanto la concepción tradicional de la evaluación se caracteriza por su énfasis en la medición de las personas. Los instrumentos típicos de esta orientación son los Tests Psicométricos. El final de este proceso evolutivo, que se sitúa en el momento actual, se caracteriza por la profundidad y extensividad científica alcanzada en el campo de la medida. Son muy numerosos los cambios que han acontecido en las últimas décadas y han supuesto un giro importante tanto en los modelos de medida, como en el campo práctico de la evaluación (Navas, 1994: 175). La medida se diversifica para atender programas, instituciones y sociedad en general.

3.1. Orígenes

Los antecedentes que dan origen a los actuales test nos ubican en las primeras pruebas sensoriales utilizadas por Galton (1822-1911). También se debe a Galton la aplicación de la estadística para analizar los datos provenientes de sus test, labor que continuará Pearson.

La influencia de Cattel (1860-1944), marca la actividad científica de la década de los noventa del siglo veinte. Es el primero en utilizar el término *Test Mental* en su artículo *"Mental Test and Measurement"*, publicado en la revista *Mind* en 1890. Será Binet quien dé un giro radical a la filosofía de los tests, al introducir en su escala tareas de carácter cognoscitivo encaminadas a evaluar aspectos como el juicio, la comprensión y el razonamiento, que según él, constituían los componentes del comportamiento inteligente. La revisión de esta escala la realiza Terman, que utiliza por primera vez el cociente intelectual (CI) para expresar las puntuaciones de los sujetos. Esta idea era originaria de Stern, quién en 1911, propuso dividir la edad mental (EM) entre la cronológica (EC), multiplicando por 100 para evitar los decimales. Estos autores constituyen referencias históricas claves en el desarrollo de escalas individuales de inteligencia.

El siguiente paso en el devenir histórico de los Test viene marcado por la aparición de test de inteligencia colectivos, propiciados por las necesidades políticas y sociales de seleccionar y clasificar soldados para tomar parte en la Primera Guerra Mundial (1917). La experiencia acumulada lleva a la conclusión de que la puntuación global que aportan los Test de Inteligencia era insuficiente para detectar

los diferentes aspectos de funcionamiento intelectual de los sujetos y se impone la evaluación de características más específicas para realizar pronósticos particulares más precisos. Ello da origen, entre los años 30 y 40, a las hoy clásicas baterías de aptitudes, que conformaban lo que entonces se consideraban los componentes fundamentales del funcionamiento inteligente: comprensión verbal, fluidez verbal, aptitud numérica, aptitud espacial, rapidez perceptiva y razonamiento general. La técnica del análisis factorial permite estructurar la genérica puntuación global de la inteligencia. Paralelamente al desarrollo de Tests Cognoscitivos, los Test de Personalidad también se benefician de los avances técnicos del análisis factorial y otras técnicas multivariadas afines. A estos avances se añaden nuevas orientaciones como los Tests Proyectivos (Muñiz, 1994).

En el marco educativo la medición surge estrechamente ligada al nacimiento de la Pedagogía Experimental, como disciplina científica, como muestran los primeros trabajos empíricos que se inician en el campo educativo y corresponden a tests de instrucción, inteligencia y escalas métricas de escritura, entre otros.

En educación la medida es problemática, en tanto la gran mayoría de variables, aspectos, factores, y dimensiones educativas no pueden medirse directamente y las explicaciones científicas se realizan a través de constructos que sirven para conceptualizar un gran número de conductas empíricas. Los constructos constituyen herramientas intelectuales imprescindibles para clasificar y describir las conductas, observar fenómenos educativos y establecer predicciones y relaciones causales. Configuran las teorías.

Del Rincón y Otros (1995: 50-51) señalan como problemas de medida en Ciencias de la Educación:

a) *Medir características no observables directamente*, lo que obliga a derivar indicadores empíricos de los constructos teóricos con la consiguiente consecuencia en la validez de constructos.
b) *El carácter cualitativo de los fenómenos educativos* que dificulta una acumulación de registros ordenados y simbolizados matemáticamente.
c) *La implicación del factor humano en todo proceso de medida*, contribuye a la inevitable presencia de sesgos y errores que se producen cuando se repite una medición.

3.2. Tendencias científicas actuales

Algunos de los desarrollos científicos actuales se concretan en: a) *Elaboración de bancos de ítems*, b) *Tests a Medida* y c) *Tests referidos al Criterio.*

a) *Los bancos de ítems* son conjuntos de elementos que miden un mismo rasgo, campo o área de conocimiento, almacenados junto con sus propiedades estadísticas, y sus parámetros. Los primeros trabajos realizados sobre bancos de ítems se realizaron a final de la década de los 60 y a comienzo de los 70, pero su uso se

hace realidad a partir de aplicaciones informáticas que facilitan el almacenamiento, la clasificación, y la estimación de los parámetros de los ítems. Los bancos de ítems contienen por tanto ítems parametrizados (calibrados en la misma escala) que definen operativamente la variable medida. Van der Linden y Eggen (1986), señalan dos ventajas fundamentales de los bancos de ítems frente a Test estandarizados. La primera de ellas tiene que ver con la flexibilidad que esta alternativa introduce en la evaluación, ya que posibilita la construcción de test basándose únicamente en consideraciones prácticas de carácter específico, relacionadas con las necesidades de evaluación que, en un momento determinado, exigen el desarrollo de un test concreto. La segunda ventaja tiene que ver con la actualización periódica de las estimaciones de los parámetros de los ítems y la posibilidad de una evaluación individualizada y referida al criterio.

Los bancos de ítems constituyen un intento de integrar la construcción y análisis de ítems con la teoría. Se han construido bancos de ítems en sectores muy distintos: fuerzas armadas, organizaciones médicas y sistema educativo.

En educación los bancos de ítems se aplican a muy distintos ámbitos educativos. Algunos se elaboran con la finalidad de evaluar el sistema educativo, tales como los bancos de ítems de la National Assessment of Educational Progress, otros son destinados a aplicaciones dentro del aula. También la evaluación de áreas curriculares es propicia para el desarrollo de bancos de ítems. Las políticas científico-educativas potencian esta vertiente científica, al apoyar líneas de trabajo en torno a la elaboración de bancos de ítems para la evaluación de sistemas educativos en base a áreas curriculares. La gran mayoría de software informático actual especializados en test incluyen la opción de creación de bancos de ítems (MicroCat, FastTest, The Examiner, etc.).

b) *Tests a Medida*. La idea de evaluación individualizada, de aplicación de test adaptados, hechos a la medida de los sujetos objeto de evaluación, no es nueva, se remonta a los trabajos de Binet y Simon sobre inteligencia en 1908. Sin embargo, apenas se trabajó sobre esta idea hasta finales de la década de los 60. El desarrollo de software informático en la evaluación individualizada potencia y consolida un nuevo área denominada CAT (Computer Adaptative Testing). A ello ha contribuido también la introducción de la TRI (Teoría de Respuesta al Item). Muñiz y Hambleton (1992), señalan como principales ventajas la adaptación al nivel de los sujetos a los que se aplican, con lo que se gana en precisión (se minimiza el error de medida) consiguiéndose asimismo efectos colaterales deseables, como el aumento de la motivación del sujeto o evitar la pérdida de tiempo en contestar ítems demasiados fáciles para su nivel. Además la longitud del test puede disminuirse manteniendo la misma fiabilidad. Estas mediciones ajustadas sólo son posibles cuando las mediciones son invariantes de los instrumentos, y este marco lo aporta la TRI. Además del problema de la invarianza, la TRI soluciona el problema central de los tests a medida que es elegir el test más adecuado a cada sujeto. El uso de la función

de información, en combinación con un banco de ítems, permite elegir aquel Test que maximice la información para los valores deseados de medida, es decir, para los valores en los que se estima que se encuentra la competencia del sujeto a evaluar. A pesar de que todos los problemas que plantea este tipo de test no están resueltos, se puede predecir, sin riesgo de equivocarnos, que en años sucesivos asistiremos a un gran crecimiento y asentamiento de los tests a medida y de evaluaciones individualizadas, justificadas a su vez por cambios en modelos y concepciones educativas.

c) *Los Tests referidos al criterio*. Uno de los avances importantes de la Teoría de los Tests en los últimos 25 años ha sido el desarrollo y creciente interés por los Tests Referidos al Criterio (TRC). Estos tests plantean procedimientos para evaluar el rendimiento de los sujetos en dominios de contenidos bien definidos, en vez de por referencia a la conducta de otros sujetos. Su uso encuentra un campo abonado en la evaluación educativa, propiciada por el énfasis que se pone en los programas e intervenciones educativas en los años sesenta. Pronto se hizo evidente que los tests tradicionales, dirigidos a la evaluación de un rasgo o aptitud, e interpretados según normas de grupo, no eran adecuados para evaluar los rendimientos educativos, ni para diagnosticar necesidades de intervención. La expresión Test Referidos al Criterio aparece por primera vez publicada en la revista *American Psychologist*, en un artículo de Glaser (1963: 519), que los define "como aquellos en los que los resultados dependen del estatus absoluto de calidad del estudiante, frente a los TNR que dependen del estatus del grupo". Esta diferenciación se establece en la cultura científica durante la última parte de los años sesenta y en los setenta. Antecedentes, no obstante, de tal diferencia los encontramos en 1913 con Thorndike, que estableció una distinción similar pero que fue olvidada por el desarrollo de la psicometría clásica. Más tarde otros autores retomaron la idea, como Flanagan, Nedelsky y Ebel, a los que no se les presta el apoyo necesario, llevados los científicos por su obsesión por la medición de rendimientos, destrezas y aptitudes según la ley de distribución normal. El eco de la aportación de Glaser se explica por el giro hacia la evaluación de programas, las intervenciones educativas, y la necesidad de establecer diagnósticos previos de los sujetos antes de la intervención, en términos de lo que eran *"capaces de hacer"* y en este sentido los tests tradicionales no proporcionaban la información necesaria. No obstante, a pesar de su interés, los TRC tuvieron un largo periodo de gestación y su implantación definitiva no se produjo hasta varios años después. Es en los años setenta cuando empieza a trabajarse de forma sistemática sobre este tipo de test, teniendo que hacer frente a críticas de teóricos de la medida aferrados a la psicometría clásica. Por tanto, es a partir de los años setenta cuando comienza a generalizarse su uso. En la década de los noventa las referencias a estos tests, en publicaciones especializadas es muy numerosa (*Journal of Educational Measurement, American Educational Research, Applied Psychological Measurement Journal*, etc.) y su aplicación educativa habitual.

El mantenido interés en la evaluación referida al criterio y el resurgimiento de la TRI han ido paralelos en el tiempo, proporcionando esta última un marco muy apropiado para la elaboración de estos tests. Así, la TRI viene a solucionar dos problemas básicos de la TRC, el primero, la unificación de la métrica para los distintos test utilizados (invarianza), y el segundo, la preparación de test más ajustados al maximizar la información que aportan. Los TRC son más apropiados para su aplicación al campo educativo y profesional, que para la medición de variables psicológicas. La proliferación de bancos de ítems nacionales e internacionales, agrupados por especialidades académico profesionales, ha supuesto un empuje para esta forma de medida. Una extensión de este planteamiento lo encontramos hoy día en la idea de Tests de Dominio Computerizado que está a caballo entre los TRC y los Test a Medida, y que se debe al trabajo de Lewis y Sheeban (1990). En este tipo de test se conjuga la idea de la evaluación individualizada con los TRC.

Una reflexión sobre el momento actual referida a los Tests y la medida en educación constituye el tema monográfico de dos recientes números (1998 y 1997) de la *Revista International of Educational Research*. En ellos se expone el panorama científico actual y su prospectiva. Uno de los temas de gran interés actual lo ocupa la medición del rendimiento escolar, que últimamente está experimentando un acelerado desarrollo y grandes cambios. La psicometría contemporánea se enfrenta a la necesidad de medir nuevos resultados cognitivos, escolares, especialmente habilidades de alto nivel cognitivo, tales como, resolución de problemas, razonamiento y pensamiento crítico. De ello se deriva a su vez, una mayor necesidad de nuevos modelos de medida, nuevas conceptualizaciones de validez y fiabilidad en los tests, y nuevos modos de medición.

La descripción del panorama actual en Test de Rendimiento Escolar puede realizarse en función de seis áreas. En cada una de ellas se identifican cambios que son consecuencia del nuevo contexto científico caracterizado por: a) la exigencia de responder a necesidades políticas y educadores, b) la emergente tecnología informática y de telecomunicaciones, y c) la disponibilidad de nuevas teorías y modelos en la psicología cognitiva, la educación, y métodos psicométricos. Por tanto, la tendencia general en la actualidad se caracteriza por:

- el incremento del uso de formatos de respuesta construida o elaborada por los sujetos, frente a las respuestas seleccionadas entre opciones múltiples;
- el incremento del énfasis por medir conductas cognitivas de alto nivel, en contraposición a habilidades cognitivas de bajo nivel;
- preferencia por los test basados en criterios;
- creciente sofisticación de los modelos de medición; incremento del rigor (validez y fiabilidad de las inferencias derivadas de los datos de los test de rendimiento);
- incorporación de las tecnologías informáticas en la elaboración, administración, puntuación, y en los informes que se emiten sobre los resultados de los test de rendimiento (Hambleton y Sireci, 1997: 379-380).

4. IMPACTO DE LAS TECNOLOGÍAS EN LAS PRÁCTICAS EVALUATIVAS

La medición educativa tecnológica se inicia como subárea científica de la medición educativa.

La revolución tecnológica, al hacerse extensiva y accesible, posibilita cambios en la evaluación, específicamente en lo que atañe a la construcción y valoración de test. Bunderson y otros (1993: 368) identifican cuatro generaciones en la evolución tecnológica en su aplicación a la medida. Cada una de ellas conlleva un cambio conceptual y cualitativo de la medida. Es precisamente aquí donde podemos mostrar la relación entre conceptos y constructos científicos y desarrollos tecnológicos, como planteamos en la introducción de este capítulo.

La denominada primera generación tecnológica de la medida se caracteriza por los Test Computerizados que implica la administración de test a través del ordenador (CT). La segunda generación la representan los Test Adaptativos Computerizados (CAT). Una revisión de diferentes aspectos psicométricos de los tests computerizados y adaptativos puede verse en Wainer (1990). La tercera generación se caracteriza por la medida continua (CM) que implica estimar la dinámica de cambio en las trayectorias de aprendizaje de los alumnos para establecer perfiles de evolución como aprendices. La cuarta generación denominada "inteligente" se caracteriza por aportar datos e interpretaciones de perfiles individuales, así como proporcionar información de los procesos de enseñanza-aprendizaje que orienten tanto a alumnos como a profesores.

Una gran mayoría de programas informáticos disponibles en el mercado se diseñan específicamente para cubrir las tareas que conlleva cualquier proceso de medición educativa: a) *elaboración de test,* incluye la redacción de test, estudio piloto y revisión de ítem, b) *administración de test* que implica obtener las respuestas, clasificar a los sujetos en función de las mismas e interpretar los resultados y c) *análisis de test* que abarca la validación del mismo y el análisis de índices de discriminación de ítem.

Estas generaciones tecnológicas, como decíamos anteriormente, van acompañadas de cambios conceptuales. Así mientras la primera y segunda generación informática asume la medida de forma estática, la segunda y tercera generación trabajan con medidas dinámicas. Es decir, en el primer caso la medición se entiende de forma estable y por tanto su objetivo es la medición de la posición de una persona o evento en un momento preciso. La medición del rendimiento escolar es un ejemplo de medida fija. En las últimas dos generaciones informáticas se entiende la medición de forma evolutiva y por tanto se registran los cambios a través del tiempo. La medida del cambio en el rendimiento individual en función del proceso instructivo es un ejemplo de medida dinámica porque el objeto de la medida es describir los progresos en el alumno a lo largo del tiempo.

Otro aspecto que marca diferencias entre estas generaciones informáticas es el cambio de objetivo educativo al que sirven. Tradicionalmente la medida ha tenido aplicaciones institucionales o individuales. Institucionalmente la medición ha servido a muy distintos propósitos: admisión de alumnos, valorar el nivel de logro de objetivos educativos, evaluar programas, o evaluar instituciones o centros. El uso tradicional de la medida para los sujetos ha sido servir a su orientación, basándose en sus habilidades, aptitudes o intereses, sobre la base de la posición de un sujeto en una escuela de medida. Frente a estos usos estandarizados de la medida, las nuevas generaciones otorgan un uso más amplio y abierto en sus aplicaciones educativas. Ayuda a establecer trayectorias de aprendizaje, diagnosticar problemas de forna inmediata o establecer perfiles y trayectorias individuales de aprendizaje.

Las condiciones en las que se toman las medidas constituye otro aspecto de cambio del que se hacen eco las respectivas generaciones informáticas. Así mientras en las dos primeras generaciones informáticas la medida se obtiene mediante situaciones estandarizadas, en las que los sujetos dan una respuesta en una determinada escala, en la tercer y cuarta generación las tareas de prueba son más complejas, teniendo cabida respuestas múltiples, más realistas y globales.

Además de estas aportaciones la informática ha posibilitado una considerable mejora de las condiciones de control de la medida. La informática en sus aplicaciones a la medida ha desarrollado básicamente cinco áreas:

1) Registro de estímulos muy variados, desde información puntual hasta dimensiones ambientales,
2) Registro y acceso a secuencias de procesos y operaciones,
3) Toma de decisiones, basadas en procesos de entrada y contabilización de información registrada,
4) Interconexión de información almacenada en distintas fuentes, formatos y lenguajes.

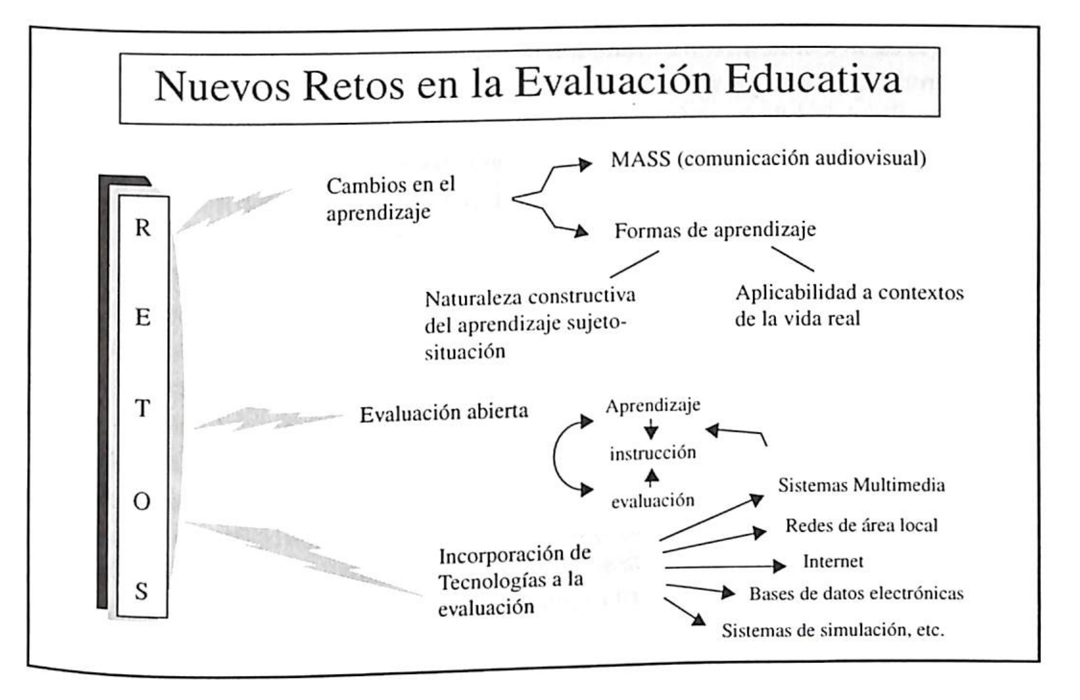
Las tecnologías de la información y la comunicación, concretamente Internet, facilita el acceso y disponibilidad de estos avances informáticos a todos los profesionales de este ámbito de trabajo, permitiendo estar puntualmente informados de todos los avances y desarrollo informáticos.

Internet pone al alcance de cualquier internauta, interesado en esta temática, software específico dedicado a la medición, estando disponible a través de Internet. Algunas referencias importantes son "Assessment Systemd Corporation: Psycometric Software"[2], este enlace proporciona y permite bajarse demos de programas específicos sobre medición; o *Epicent*[3], dirección que posibilita localizar compañías que trabajan con software dedicado a la medición y a los test. También editoriales de gran prestigio académico abren una nueva línea de servicios de software. Otras direcciones[4] permiten el acceso a base de datos especializadas en temáticas de la medida.

[2] http:// www.assess.com
[3] http://www.epicent.com/software/subjets/measurement/measurement.html

5. PERSPECTIVAS Y NUEVOS RETOS EN LA EVALUACIÓN EDUCATIVA

Los cambios educativos, provocados por los cambios sociales y culturales, plantean nuevos retos a la medida y la evaluación educativa.



Entre ellos podemos mencionar los cambios que afectan a los aprendizajes, consecuencia de la implantación de los MASS (comunicación audiovisual), y la inclusión de diferentes formas de aprendizaje; aprendizaje basado en la resolución de problemas, aprendizaje orientado a proyectos, o aprendizaje a distancia. Dos son las proyecciones más destacables de estos cambios: la naturaleza constructiva del aprendizaje, y la ampliación del aprendizaje al contexto de la vida real. La primera está avalada por una línea de investigación psicológica que plantea el aprendizaje como una actividad y un proceso constructivo que se realiza por medio de la interacción con sus ambientes. Estos planteamientos ofrecen un marco conceptual para prácticas educativas, y lleva a reconsiderar los instrumentos de evalua-

[4] http://www.sagepub.com/sagepage/software.htm

ción en el sentido de que la acumulación de aprendizaje en un mismo dominio y bajo determinadas circunstancias, no es un punto de vista apropiado con relación a los objetivos que hoy día la educación se plantea como más importantes (por ejemplo, la comprensión y resolución de problemas).

Tampoco resultan apropiados en formas de aprendizaje que implican la construcción de significados por parte de los alumnos, y el desarrollo de estrategias para abordar nuevos problemas y tareas de aprendizaje. Por tanto se requiere de nuevos tipos de instrumentos que permitan evaluar niveles cualitativamente distintos de comprensión, así como estrategias diferentes en el abordaje de problemas no familiares y cambios en las situaciones de aprendizaje. Por tanto se necesita de un nuevo tipo de instrumento de evaluación desde la perspectiva de una mayor integración entre instrucción y evaluación (Dochy, 1992).

En el campo de la práctica se cuenta con algunas propuestas tales como el Test de la Universidad de Maastricht (Segers, 1996), OAT (Over All Test), que se focaliza sobre todo lo que el alumno plantea para analizar los problemas y contribuir a sus soluciones con la aplicación de conocimientos adquiridos, modelos y teorías. Este test mide si el estudiante dispone de los conocimientos relevantes que necesita para resolver los problemas que se le presentan. Esta idea también conlleva el entrelazamiento o multidimensionalidad de factores de aprendizaje; conocimiento, destrezas, capacidad de relación, selección de aspectos relevantes, etc.

Las perspectivas futuras se observan como integración del aprendizaje, la instrucción y la evaluación a través de ambientes de aprendizaje flexibles y transformadores. "*La evaluación abierta*" es un concepto que planea en esta visión de futuro. Ello implica cambios conceptuales importantes, tales como invertir el concepto de evaluación, que pasa de concebirse como un medio para determinar lo que los individuos han conseguido y la medida de su adaptación a los medios de instrucción, a adaptar las condiciones de instrucción a los individuos para maximizar su potencial de éxito. Este objetivo sólo puede ser realizable si el aprendizaje puede ser diseñado teniendo en cuenta el perfil individual de conocimientos y habilidades (Dochy y Moerkerke, 1997).

Junto a los cambios sobre evaluación también se observan cambios importantes en los instrumentos para llevarla a cabo. En este sentido, adquieren protagonismo los sistemas multimedia, las redes de área local, los sistemas de telecomunicación como Internet, bases de datos electrónicas, materiales electrónicos de autoestudio, sistemas de monitorización a distancia y en red, etc. También el audio constituye un soporte importante para la evaluación de dimensiones formativas difíciles de observar a través de formas escritas. La informática posibilita también la disponibilidad de complejos sistemas de simulación que permiten la expresión de complejas habilidades prácticas.

Otra idea que parece tener cierta proyección concierne al valor de la evaluación para el propio aprendizaje del alumno. La evaluación se convierte en herra-

mienta para diagnosticar las condiciones iniciales de su aprendizaje y establecer metas de logro. La evaluación permite revisar y regular en cada momento estas dos facetas tan importantes en el aprendizaje.

Por tanto podemos concluir que el campo de la medición educativa y la evaluación escolar ha vivido un progreso importante en los últimos años. Las contribuciones de las teorías de la inteligencia, de la psicometría y del escalamiento, no pueden ser infravaloradas, pero también hay que considerar el impacto de los cambios en la sociedad que afectan muy directamente a las prácticas evaluativas. Entre estos podemos mencionar el rapidísimo avance de las nuevas tecnologías de la comunicación, que a su vez propician poderosos ambientes de aprendizaje, y generan grupos muy heterogéneos de personas que incrementan su demanda educativa en formas más flexibles. La relación de poder entre alumnos y profesores cambia. Los alumnos asumen mayor responsabilidad y tienen mayor capacidad de decisión en lo que concierne a lo que ellos quieran aprender y cómo lo quieran aprender.

Por otra parte, la naturaleza y el almacenamiento de conocimientos y habilidades es diferente. Todos estos cambios están afectando a la evaluación. De ahí que en un futuro inmediato la evaluación adopte cambios relacionados con:

a) La importancia del lugar o contexto a lo largo de todo el proceso de aprendizaje.
b) La focalización tanto sobre el conocimiento adquirido, como sobre la adquisición de habilidades y competencias.
c) La adopción de diferentes tipos de perfiles o estilos de aprendizaje.
d) Uso extensivo los formatos educativos que ofrecen la tecnología de las comunicaciones y sistemas multimedia.
e) Diseños de evaluación diversificados y dirigido a muy distintas funciones.

Dentro de las futuras funciones que la evaluación puede cubrir, se plantea que la nueva cultura evaluativa puede ser utilizada para hacer cambiar el concepto instruccional desde una visión depositaria del conocimiento, a un intento de desarrollar en los estudiantes la idea de capacitarse para "aprender a aprender". Existen algunas experiencias en este sentido como la que llevan a cabo Clark y Stephens (1996), que utilizan la evaluación como medio para una reforma sistemática de las matemáticas en Australia.

ACTIVIDADES PARA EL ALUMNO

1. Idea una representación gráfica (dibujo, ideograma, infografía, cómic, etc.) que visualice los aspectos que constituyen una problemática específica de la medida en educación.
2. Elabora un mapa conceptual propio que represente tu visión del panorama actual de la evaluación educativa desde un punto de vista científico. En él han de quedar reflejadas las principales áreas de investigación.

3. Haz un viaje imaginario a través del tiempo y sitúate en el futuro. Describe que rasgos o características tendrá la evaluación educativa. Razona tu respuesta.
4. Consulta alguna de las siguientes direcciones electrónicas e indica algunas de sus principales contenidos y /o actividades.

http://www.assess.com

http://www.epicent.con/software/subjects/measurement/measurement.html

http://www.sagepub.com/sagepage/software.htm

5. Consulta algunos índices de las revistas especializadas recomendadas e indica temas que se abordan en los últimos números.

BIBLIOGRAFÍA COMENTADA

Textos y Manuales

BUENDÍA EISMAN L. (Coord.) (1996). "Evaluación educativa" (Monográfico). *Revista de Investigación Educativa,* Vol. 14, nº 2.

En este monográfico se recogen aportaciones que cubren un abanico amplio de temáticas sobre evaluación: investigación evaluativa, evaluación de alumnos universitarios, evaluación del profesorado, elaboración de pruebas para la evaluación etc.

DEL RINCÓN, D. Y OTROS (1995). *Técnicas de investigación en Ciencias Sociales*. Madrid: Dykinson.

Este texto presenta una panorámica amplia de técnicas frecuentemente utilizadas en la evaluación. Se exponen como técnicas para obtener información los Tests, pruebas objetivas, escalas de medida, cuestionarios, observación sistemática y participante, entrevistas y documentos. Se describe el proceso de elaboración y construcción de cada una de estas técnicas. También estos autores abordan la validez y fiabilidad de los Test.

DE PABLOS, J. (Coord.) (1998). *Nuevas Tecnologías*. Barcelona: CEDECS.

Esta obra pretende profundizar en la compleja realidad que implican las nuevas tecnologías, invitando al lector a la reflexión sobre la trascendencia e impacto de las tecnologías aplicadas en sentido global y en la educación de forma concreta. Profesores e investigadores procedentes de muy distintas universidades y áreas de conocimiento colaboran en este libro.

Revistas Especializadas Internacionales

Applied Measurement in Education (AME)

Esta publicación está dedicada a la investigación sobre aplicaciones de la medición educativa y psicológica, incluyendo investigaciones aplicadas centradas

en los usos de medidas en contextos educativos, estrategias novedosas para resolver problemas de medición educativa, revisiones de investigaciones pertenecientes a usos contemporáneos de la medición, y revisiones comparadas de test y de métodos utilizados en mediciones educativas específicas.

Educational Assessment

Se publican cuatro números al año, la revista contiene una selección de investigaciones y estudios relacionados con pruebas y mediciones educativas. La revista presenta también estudios originales sobre aprendizaje y conocimiento con el objeto de proporcionar un foro a expertos interesados en la proyección práctica.

Educational and Psychological Measurement (EPM)

Esta revista recoge trabajos que abordan problemas de la medición de diferencias individuales. Los artículos versan sobre desarrollos y usos de test y medidas en educación, organizaciones e instituciones; descripciones de programas de medición (testing) con múltiples fines, incluyendo además sugerencias de nuevos tipos de ítems o métodos de tratamiento de datos de test.

Journal of Educational Measurement (JEM)

Esta revista publica trabajos de teoría y práctica de medición educativa. Como representación citar exposiciones de prácticas de medición educativa, informes de nuevos usos de la medición en educación u otras disciplinas relacionadas, argumentaciones filosóficas sobre la medida, procedimientos para informar e interpretar resultados de la medición, e investigaciones sobre proceso, técnicas, herramientas y procedimientos relacionados con la medida.

Wilson Mark, E. *Objective measurement: Theory into practice.* University of California: Berkeley. Ablex Publishing Corp. At Press L.T.D.

Los volúmenes de esta serie que comienza a plublicarse en 1992 (v.1), presentan investigaciones originales sobre aplicaciones, teorías, filosofía e historia de la medida. Recogen las aportaciones de las conferencias bianuales de *"The International Objective Measurement Workshop"*. Las contribuciones se agrupan en torno a tres temáticas: prácticas de la medida, teoría de la medida y aplicaciones matemáticas y estadísticas para la medida.

CAPÍTULO II

EVALUACIÓN Y GESTIÓN DE LA CALIDAD EDUCATIVA

Teresa González Ramírez[1]
Universidad de Sevilla

OBJETIVOS

1. Conocer la importancia en la actualidad de la calidad para la educación
2. Presentar distintos conceptos y definiciones sobre el concepto de calidad en educación
3. Introducir el concepto de Gestión de la Calidad Educativa dentro del movimiento de la Calidad Total
4. Analizar las herramientas para la Gestión de la Calidad Educativa
5. Presentar las Normas de Garantía de Calidad aplicadas a contextos formativos y/o educativos

CONTENIDOS

---

[1] Facultad de Ciencias de la Educación. Área de Métodos de Investigación y Diagnóstico en Educación. Av. San Francisco Javier s/n, 41005-SEVILLA. Tfno: 954556837. E-mail: tgonzale@cica.es

INTRODUCCIÓN

Vivimos en una sociedad que en todas sus manifestaciones demanda calidad. Este concepto puede ser analizado desde muchas perspectivas, enfoques y modelos; su recorrido en el tiempo nos muestra como desde una concepción etimológica con una valor absoluto ha pasado a considerarse como concepto emergente, contextual, polivalente e integral.

Las razones de esta evolución no son producto de la casualidad; los adjetivos que hemos utilizado anteriormente aplicados al concepto de calidad son extrapolables a cualquier ámbito de conocimiento e investigación; desde una concepción epistemológica sistémica (De Miguel, 1997) con un marco de investigación fundamentado en el realismo científico (Pawson y Tilley, 1997), los "contextos de la evaluación" como dice la profesora Colás en el capítulo anterior, se han diversificado y en este panorama científico y metodológico el concepto de calidad no es más que una consecuencia de ese devenir histórico.

Este capítulo se fundamenta desde una concepción de la educación con una visión organizacional e institucional; iniciamos su desarrollo a partir de la importancia y el concepto de calidad educativa en la actualidad como una consecuencia de la dinámica que han sufrido las estructuras educativas y/o formativas. Desde esta visión justificamos el concepto de *gestión de la calidad educativa* como un nuevo constructo que emerge en el contexto educativo más ajustado a los nuevos tiempos; supone una superación en el tiempo del concepto de evaluación; factores (de índole económico, político, científico y cultural), en la década de los noventa han impulsado como dice Rodríguez Espinar (1998:134) que "la evaluación se constituya como una fase naturalmente necesaria en cualquier proceso de intervención, sin evaluación no es posible progresar racionalmente", de manera paralela, también ha propiciado que la evaluación se convierta en la herramienta para gestionar la calidad de la educación con objeto de tomar decisiones de manera razonable y fundamentada, motivadas a veces por instancias externas, otras de carácter interno, como consecuencia de la propia evolución natural de la institución, pero en definitiva, el marco social en el que nos movemos está pidiendo que optimicemos los procesos y mejoremos los resultados de la educación.

Desde esta perspectiva, es necesario conocer los principios y características de la gestión de la calidad y las herramientas de la calidad para proyectarlas en los contextos educativos; aunque su origen se remonta al ámbito del sector productivo su presencia se está haciendo eco en cualquier tipo de organización (una muestra de ello es el esfuerzo que está haciendo la Asociación Española de Normalización y Certificación en revisar las normas de calidad para que se puedan ajustar a cualquier contexto organizativo). En este sentido, es necesario hacer una llamada sobre lo que hemos denominado la "*ética de la responsabilidad*"; implantar un sistema de calidad en una institución educativa tiene dos aspectos diferenciados: un aspecto

técnico "saber hacer" y otro humano "querer hacer"; este último cobra sentido desde la ética de la profesión docente.

Termina este capítulo analizando la importancia que ha tenido la política de gestión de la calidad en nuestro país en todos estos planteamientos. El hecho de que se hayan tomado medidas ministeriales para gestionar la calidad de los centros educativos públicos, que el Consejo de Universidades elabore un Plan para desarrollar procesos evaluativos que sirvan para gestionar la calidad de la institución universitaria, o que el Consejo de la Unión Europea invite a sus estados miembros a adoptar mecanismos de evaluación de la calidad y garantía de la calidad para los sistemas de enseñanza superior, o se creen organismos específicos como es el Instituto Nacional de Calidad y Evaluación, no es más que un exponente del papel que juegan las políticas educativas como factores externos a la organización a la vez que sirven de impulso para pulsar y modular las exigencias que la sociedad en general demanda a la educación. La evaluación desde esta perspectiva reviste una gran importancia estratégica y está al servicio del cumplimiento de las finalidades del sistema y del conjunto de todos los componentes que están al servicio del desarrollo de la calidad.

1. LA IMPORTANCIA DE LA CALIDAD PARA LA EDUCACIÓN

En la última década, los temas de Calidad, Garantía de la Calidad y Gestión de la Calidad se han convertido en preocupación y objetivo prioritario para las organizaciones. En sus inicios, fue el mundo industrial el primero en reconocer que gran parte del crecimiento y éxito económico en Japón podrían deberse a la insistencia de las empresas de ese país en la calidad y el perfeccionamiento de la calidad. Fue sobre todo a partir de la década de los ochenta cuando en el contexto europeo se empezaron a adoptar, los principios y métodos de la Gestión de la Calidad Total para cumplir el nivel de calidad deseado por sus clientes y para mejorar continuamente la calidad de los productos y servicios que ofrecen. Uno de los rasgos más visibles en Europa de esta "tendencia hacia la calidad" ha sido la certificación de los mecanismos de garantía de calidad en base a las normas de calidad.

En este contexto, no es sorprendente que la educación haya mostrado también su interés por este tema. La formación profesional y la formación continua, en el contexto empresarial, fueron los primeros sectores interesados por conocer las normas de calidad y los principios de la calidad. Ambos tipos de organizaciones se hallan de hecho más próximos a las necesidades del mercado y a la evolución económica que las instituciones educativas regladas; sin embargo esta evolución está dando sus primeros pasos y tanto la cultura de la calidad como las herramientas de la calidad están empezando a ser exigencias ineludibles en los contextos educativos y/o formativos.

Por otra parte, también hay que decir que este interés por la calidad no es nuevo en el ámbito educativo: instituciones, profesores, alumnos, administradores

y decisores políticos siempre han estado interesados por este tema; esa preocupación se ha extendido hacia la utilización de normas y procedimientos que garantizaran la calidad de la enseñanza impartida; aunque también tradicionalmente, la calidad se ha interpretado de forma muy estrecha, centrándose únicamente en características particulares de los servicios educativos o formativos ofrecidos.

En este sentido, el concepto de calidad ha evolucionado con el tiempo. La tendencia actual de la calidad es centrarse cada vez mas en la eficacia global de un ofertor educativo, ya sea una escuela profesional, una universidad o una institución formativa privada. Esta tendencia corre de forma paralela a la propia evolución industrial, en la que las consideraciones de calidad se desplazan hacia la capacidad organizativa de ofrecer productos y servicios de alto nivel de calidad.

Los nuevos enfoques de garantía y gestión de la calidad educativa implican a grandes rasgos (Van den Berghe, 1998):

a) Un cambio en los esfuerzos de calidad en las escuelas, que pasan de centrarse en la calidad del docente hacia el estudio y optimización de la institución en su conjunto.
b) La aplicación de mecanismos de control de la calidad nuevos o complementarios en la enseñanza superior.
c) La creación por primera vez de sistemas de garantía de la calidad y mecanismos relacionados con el rendimiento dentro de la formación continua.

1.1. El énfasis en la calidad como resultado de la dinámica de las estructuras educativas y formativas

Los actuales sistemas educativos deben su configuración a la confluencia de factores de diversa índole: social, económica, pedagógica, profesional y cultural. Todos estos factores siempre han existido pero su importancia relativa ha ido cambiando a lo largo del tiempo de tal manera, que en la actualidad, las demandas exteriores de los sistemas educativos (desde los gobiernos, estudiantes, empresarios, etc...) se incrementan cada vez más, presionando sobre el desarrollo de los recursos y la eficacia de la organización. Por otra parte, el sector de la formación continua se está convirtiendo en un sector económico consolidado y maduro dentro del sector productivo "servicios".

Estas tendencias nos apuntan que la educación ahora mas que nunca está empezando a considerarse como un sector económico normal. Esto significa que escuelas, universidades y formadores en general deben ofrecer un rendimiento mayor, comportarse de manera más profesional y ofrecer permanentemente servicios de calidad. Compartimos con Van den Berghe (1998: 8) en que "los paradigmas de la educación están cambiando desde una enseñanza dirigida por la oferta, hacía un aprendizaje dirigido por la demanda".

Entre los factores específicos que están contribuyendo a este creciente énfasis sobre la calidad en la educación podemos citar los siguientes:

a) *La calidad se ha convertido en una exigencia de la sociedad actual.* Cuando se habla de calidad, la amplia gama y altos niveles de calidad de productos y servicios que presentan los países desarrollados hacen aumentar las exigencias de los ciudadanos e incrementan su espíritu crítico hacía el rendimiento de baja calidad. De este espíritu de exigencias también se ha hecho partícipe la educación.
b) *La calidad es un factor de cambio, flexibilidad y personalización.* La sociedad en la que vivimos es polivalente, las formas de vida y las expectativas de las personas son menos uniformes que hace unos años. En educación esta tendencia se aprecia a través de los requisitos de cualificación cada vez más variables y complejos.
c) *La calidad nos lleva a la calidad.* Cuanto más conocimiento se genere e información se disponga sobre la calidad de los sistemas educativos, mayor demanda se hará de la misma.
d) *La calidad supone compromiso.* Las instituciones educativas son cada vez más responsables públicamente de lo que hacen y de demostrar que ofrecen un servicio de calidad.
e) *La calidad como medio de supervivencia.* Un sistema educativo bien subvencionado sólo podrá subsistir dentro de una economía próspera. En el entorno actual toda ineficacia o carencia de flexibilidad se penalizará con menos recursos destinados a la educación.
f) *La calidad implica a muchos agentes.* En contextos históricos anteriores, la calidad de la educación podía atribuirse casi de manera exclusiva a las capacidades inherentes de los profesores. Debido a la rápida evolución del contexto histórico, estas capacidades individuales ya no bastan para garantizar un nivel de calidad.
g) *La imagen de una institución educativa es la de su calidad.* El prestigio y valía de una institución educativa ya no viene avalado por sus características o reputación histórica sino por la calidad real de una institución particular. En este contexto resulta necesario, mantener una imagen de alta calidad para los usuarios y ser capaz de demostrar su calidad de manera permanente.
h) *La calidad nos permite conocer los resultados de las acciones formativas emprendidas.* Para conseguir una educación de calidad es necesario invertir en formación y recursos; de forma paralela es necesario conocer el resultado de los presupuestos formativos en pro de la calidad.
i) *La calidad obliga a la transparencia.* La variabilidad y complejidad cada vez mayor de propuestas y planes formativos hace necesario crear mecanismos que posibiliten la transparencia de la calidad ofrecida.
j) *La calidad en su estado final se proyecta en una cultura de la calidad.* La implantación de un sistema de calidad en una institución educativa cobra sentido con el cambio de actitudes dentro de la propia institución.

Estos factores exigen en su conjunto una mayor atención hacía la calidad de la educación y la elaboración de mecanismos, procedimientos y sistemas que puedan ayudar a garantizar continuamente la calidad. Algunos de esos factores son internos a las estructuras educativas, aunque la mayoría son exteriores a ellas. Los sistemas educativos están cada vez más en estrecha interdependencia con el resto de la sociedad y están por ello sometidos a las presiones y tendencias correspondientes.

2. APROXIMACIÓN AL CONCEPTO DE CALIDAD EDUCATIVA

El concepto de calidad educativa es multidimensional y no todos los autores utilizan el término en el mismo sentido. En este epígrafe vamos a enfrentarnos con la tarea de hacerlo, pese a la dificultad que ello comporta, pero creemos de su conveniencia ya que como dice Santos Guerra (1991), es un término que con demasiada frecuencia se presenta con pretendida univocidad.

Desde un punto de vista etimológico el concepto calidad tiene dos significados fundamentales: entendido como cualidad, se identifica como conjunto de atributos o propiedades referidas a algo o alguien; en segundo lugar también hace referencia a la calidad como superioridad o excelencia, como grado que expresa la bondad de una cosa.

Coincidimos con De Miguel (1997) en que el concepto de calidad es relativo y contextual ya que permite ser definido desde múltiples perspectivas, por distintas audiencias con intereses claramente diferentes y en situaciones contextuales muy diversas. Analizando definiciones que van mas allá del plano genérico y que se sitúan en un nivel mayor de especialización vamos a presentar distintas aproximaciones al concepto de calidad en la que se priorizan determinados elementos sobre otros.

Siguiendo a Bernillón y Cerutti (1989) la calidad consiste en: hacer bien el trabajo desde el principio, responder a las necesidades de los usuarios, administrar óptimamente, actuar con coherencia, un proceso o modo de hacer; satisfacer al cliente/usuario; disfrutar con el trabajo y ofrecer lo mejor de uno mismo; reducir costes inútiles, evitar fallos; ser más eficaz, eficiente, productivo,...

También hay otras definiciones que relacionan las características de un producto con el objetivo propuesto, que combinan la existencia de ciertos atributos innatos y la satisfacción del usuario, integrando todo ello bajo la exigencia de la aptitud para su utilización. De acuerdo a este criterio la calidad se define como (Salmerón, 1988):

> *"grado de coincidencia con las exigencias, como la concordia entre las propiedades de un producto, elemento material y el fin a que se destinan".*

La mayoría de las definiciones hacen referencia a la calidad como satisfacción de necesidades. Este el caso de la definición dada por la International Standard Organization (ISO) y Aenor entre otras:

> *"Conjunto de propiedades y características de un producto o servicio que le confieren la aptitud de satisfacer necesidades explícitas o implícitas".*

López Rupérez (1994, pp. 11) cuando se refiere al concepto de calidad lo realiza en los siguientes términos:

> *"Como si de una palabra mágica se tratara, el término calidad evoca en la mente de las personas la referencia a un valor seguro, se trata de un atributo o conjunto de atributos de los objetos, de los servicios, o de las relaciones que circulan en el seno de las sociedades modernas y que, según la percepción del ciudadano, satisfacen sus expectativas razonables haciéndolos dignos de confianza".*

Por su parte, De Miguel y otros (1994: 15), ofrecen una clasificación en la que consideran la calidad educativa como un concepto multidimensional, que puede ser operativizado en función de variables muy diversas. En la tabla de la página siguiente se relacionan algunas de las opciones frecuentemente utilizadas según Garvin (1984) y Harvey y Green (1993).

Finalmente, también Rul (1995: 295, 318) propone su propia clasificación. Señala que el concepto de calidad es un mito de las determinadas políticas. Considera que se emplea en educación como justificación de determinadas políticas. Considera que toda aproximación al concepto ha de representar una necesaria desmitificación. Para él, los planteamientos sobre la calidad se abordan con demasiada frecuencia de una manera reduccionista, sin ir más allá de unos enunciados formales y de afirmaciones vacías de contenidos efectivos. Coincide que el concepto de calidad es claramente dinámico y transformacional. Concretamente según este autor, el concepto de calidad aplicado al ámbito educativo, se puede manifestar en dos perspectivas fundamentales:

a) En primer lugar la *perspectiva axiológica, referida al potencial humano.* Se deriva del concepto griego de "areté", en el sentido de excelencia, perfección, valor. Significa por tanto excelencia y altura de miras en el cumplimiento o en el desarrollo de las funciones. Es una orientación modulada por el potencial de la experiencia y la sabiduría y se relaciona con el concepto griego de "sofía", en el sentido de experiencia y perspicacia. Así pues *calidad es la tendencia a la excelencia fundamentada en el potencial de la experiencia y sabiduría de las personas.*

b) La perspectiva del mundo de la vida, referida a la capacidad de transformar información en acción. Se trata de la capacidad de gestión de una persona o de una organización determinada que transforma las ideas y los recursos en realidades beneficiosas. Se vincula con el término griego "tecné", entendido como arte, creación. En este sentido, *la calidad es la*

1. *CALIDAD COMO EXCEPCIÓN* a) Calidad como algo especial, distingue unas instituciones de otras a pesar de que es difícil definirla de forma precisa. b) Visión clásica: distinción, clase alta, exclusividad. c) Visión actual: La EXCELENCIA – Excelencia en relación con estándares. Reputación de los centros en función de sus medios y recursos. – Excelencia basada en el control científico sobre los productos según unos criterios: "centros que obtienen buenos resultados".
2. *CALIDAD COMO PERFECCIONAMIENTO O MÉRITO* a) Calidad como consistencia de las cosas bien hechas, es decir que responden a los requisitos exigidos: "Centros donde las cosas se hacen bien". b) Centros que promueven la cultura de la calidad para que sus resultados sean cada vez mejor evaluados de acuerdo con criterios de control de calidad.
3. *CALIDAD COMO ADECUACIÓN A PROPÓSITOS* a) Se parte de una definición funcional sobre la calidad, lo que es adecuado o bueno para algo o alguien: – Centros donde existe una adecuación entre los resultados y los fines u objetivos propuestos. – Centros donde los programas y servicios responden a las necesidades de los clientes. b) Centros que cubren satisfactoriamente los objetivos establecidos en el marco legal.
4. *CALIDAD COMO PRODUCTO ECONÓMICO* a) Aproximación al concepto de calidad desde la perspectiva del precio que supone su obtención: – Centros eficientes al relacionar costos y resultados. – Centros orientados hacía la rendición de cuentas.
5. *CALIDAD COMO TRANSFORMACIÓN Y CAMBIO* a) Definición de calidad centrada sobre la evaluación y la mejora a nivel institucional: – Centros preocupados por mejorar el rendimiento de los alumnos e incrementar el valor añadido. – Centros orientados hacía el desarrollo cualitativo de la organización (desarrollo organizacional).

Tabla 1. El concepto de Calidad.

habilidad ingeniosa o artística que sabe incidir en una realidad dada transformándola mediante la guía de la experiencia y del conocimiento.

Estas consideraciones sobre el concepto de calidad desde la perspectiva de diferentes autores nos lleva al establecimiento de unas características básicas del concepto:

a) *La calidad tiene un valor contextual*. Al estar relacionada con parámetros temporales, ideológicos y económicos es imposible establecer una única definición del concepto de calidad. El concepto de calidad educativa está en estrecha interdependencia con el contexto en el que se utiliza.

b) *La calidad ha de valorarse de forma plural*. No es un concepto monolítico. Se puede interpretar en relación al logro de unos objetivos, nivel de satisfacción del usuario, la excelencia en procesos y resultados, desarrollo adecuado de habilidades y aptitudes, como eficiencia, como la realización de sí mismo, etc...

c) *Tiene un carácter diacrónico*. Según la OCDE (1991) conseguir la calidad puede suponer un alejamiento radical e inmediato de las disposiciones y prácticas establecidas y plantearnos en cada momento los objetivos de la sociedad y los propósitos de la educación valorados institucionalmente.

d) *La calidad de la educación no es una actividad meramente tecnocrática*. Es un proceso impregnado de valores. La conducta de calidad no se produce automáticamente al existir unas exigencias y saber qué cosas son valiosas en el ámbito de esas exigencias. "Las exigencias unidas a las cosas valiosas se convierten en motivos; además hay que elegir cuáles son las valiosas y priorizarlas". Esta característica intrínseca a la calidad educativa nos lleva al problema de la sustantividad política: para que la calidad educativa sea posible es necesario políticamente creer que la educación es un bien social.

A continuación, vamos a concretar el concepto de calidad atendiendo a dos criterios fundamentales: a) según si se centran en el proceso o en el producto y b) según en el ámbito de la calidad en el que se fundamenten

2.1. La calidad según la prevalencia del proceso o producto educativo

El contexto del que emana el concepto de calidad ha desencadenado que la mayor parte de las definiciones relacionan la calidad con los resultados. Según Cobo (1985:358) "una educación será de calidad en la medida en que todos los elementos que intervienen en ella se orientan a la mejor consecución posible".

Para De la Orden (1993) la calidad educativa se identifica con un producto educativo válido; entendiendo la validez como funcionalidad, como congruencia con las necesidades, expectativas y aspiraciones educativas de los individuos y la comunidad y que en la práctica resulta también eficaz.

Otra dificultad intrínseca resulta de las distintas acepciones que puede tener el producto educativo (actitudes positivas hacía el aprendizaje, adquirir e integrar conocimientos, extender y refinar el conocimiento, utilizar conocimientos significativos... etc.). A menudo las definiciones se centran en los resultados, equiparando éste al rendimiento de los alumnos. Sin embargo, centrarnos únicamente en la dimensión alumnos tiene el inconveniente de equiparar la mejora de la calidad con resultados académicos de los escolares. En este sentido afirma Tiana (1993: 290):

"La evaluación, entendida como una simple revisión de los productos finales hace perder de vista la consideración del centro como un ecosistema que permite explicar y dar sentido al funcionamiento general y a los procesos que desarrollan la actividad. De la misma manera la consideración de una de las partes pierde sentido aislado de la referencia al todo y de su vinculación con el contexto".

De esta concepción se desprende que un movimiento hacia la calidad lo que ha de buscar es mejorar el proceso que busca los resultados.

Desde una perspectiva de la calidad centrada en el proceso, Esteban y Montiel (1990: 75) ofrecen una definición de calidad que ofrece una concepción muy aproximada a la que queremos expresar en este trabajo. Así dicen:

"La calidad entendida como proceso se define como principio de actuación que no apunta exclusivamente a la obtención de unos resultados inmediatos/finales, sino, fundamentalmente, a un modo de ir haciendo, poco a poco, las cosas para alcanzar los mejores resultados posibles en orden a lo que se nos demanda y a las posibilidades y limitaciones reales que se tienen."

Las definiciones de calidad aportadas según se centren en el proceso o en el producto, suponen una limitación superada ya en el tiempo. Actualmente, desde una educación de calidad de carácter multidimensional las relaciones entre distintos elementos del contexto, el proceso y el producto sería una característica común de las diversas manifestaciones educativas de calidad. En este sentido apoyamos el *concepto de coherencia* de Gil (1988) entendido como interacción y sentido de vinculación de todos aquellos elementos que constituyan el objeto de evaluación (programa educativo, centro escolar, sistema educativo, etc.) y sus indicadores. Siguiendo a este autor, "El factor último que determina la calidad educativa de un centro no son los indicadores de calidad individualmente considerados, sino la estabilidad de sentido de ilación que exista entre ellos".

Por lo tanto, la calidad de la educación vendría dada por la coherencia de cada componente con todos los demás y se caracteriza por la interrelación entre tres elementos: eficacia, eficiencia y funcionalidad.

2.2. La calidad según el ámbito educativo en el que se fundamenta

Los factores que tienen mayor incidencia en la calidad educativa según un estudio empírico realizado por Cano García (1998) de carácter bibliométrico son: Profesorado, Currículum, Evaluación y Organización Escolar. Esta categorización cruzada con variables de contexto, proceso o producto da como resultado la tabla que mostramos a continuación:

	Entrada	*Proceso*	*Producto*
– *Aspectos curriculares*		X	
– *Aspectos de organización del centro*		X	
– *Aspectos relativos al alumnado*	X		X
– *Aspectos relativos a los recursos*	X		
– *Aspectos relativos a la evaluación*			X
– *Aspectos relativos al profesorado*	X	X	

Tabla 2. Factores que inciden en la calidad educativa en relación con variables de contexto, proceso y producto (Cano García, 1998: 87).

A partir de esta clasificación, vamos a analizar distintas definiciones de calidad en cada uno de esos aspectos. Sólo nos centraremos en aquellos que por el interés y conexión que tienen con la segunda parte del contenido de esta publicación resulten de interés para conectar los distintos ámbitos de gestión de la calidad educativa tratados.

- ***La calidad centrada en el profesorado:*** existen afirmaciones que hacen depender exclusivamente la "garantía de calidad" de la formación de su profesorado; desde esta perspectiva, la responsabilidad del sistema educativo se descarga en ellos y consideran que la mejora de la calidad del profesorado pasa por: procedimientos rigurosos de selección que permitan elegir solamente a los candidatos más cualificados y muy motivados; una formación pedagógica inicial breve y basada en la práctica; una remuneración suficientemente motivante que evite la fuga de estos profesionales a otras profesiones; planes de desarrollo profesional y una administración que ofrezca posibilidades de promoción. Otros elementos que han sido considerados importantes son la estabilidad del profesorado; el trabajo en equipo y la toma de decisiones compartida; la planificación y coordinación curricular a la vez que un alto grado de autonomía. Desde esta perspectiva como señala Elliot (1993) la evaluación del profesorado puede interpretarse como otra estrategia más del poder coercitivo. Desde un modelo comprensivo la calidad educativa centrada en el profesorado debe atender tanto a la mejora del profesor como a la de la institución.

  Otro aspecto relacionado con la calidad del profesorado se centra en el reconocimiento social, profesional y económico. El documento "las cifras claves en la Unión Europea" según informe de la Comisión Europea (1996) analiza la profesión docente en la Unión Europea: los resultados más importantes se sitúan en relación al incremento del numero de años necesarios para la formación docente, las diferencias retributivas entre los distintos niveles educativos tienden a igualarse y se reducen las divergen-

cias entre salarios mínimos y máximos. De todos los aspectos señalados la formación del profesorado se considera como de los más influyentes, sin embargo debemos considerar no sólo si existen opciones de formación permanente, sino en qué modalidades se estructuran, en qué horarios se ofrecen y qué consecuencias tiene la formación impartida sobre la calidad de la docencia.

- ***La calidad centrada en los programas formativos:*** desde este ámbito de la calidad la evaluación del currículum y de los procesos de enseñanza-aprendizaje serán objetivos fundamentales. La calidad del currículum consiste en planificar, proporcionar, y evaluar el currículum óptimo (según los criterios de cada país) para cada alumno, en el contexto de una diversidad de individuos que aprenden (Wilson, 1992: 34). La calidad de los procesos de enseñanza-aprendizaje viene garantizada cuando posibilita, potencia y produce el resultado de humanizar más a todos y a cada uno de los agentes implicados.
- ***La calidad centrada en centros/instituciones:*** la calidad de los centros e instituciones educativas no podemos lograrla a través únicamente de medidas políticas diseñadas de manera estandarizada y uniforme. Nuestra historia más reciente ha mostrado que si bien estas decisiones han servido de impulso y cambio, se han vuelto incapaces de asegurar la calidad educativa de manera equitativa en las diversas escuelas del país. Los estudios de casos realizados por la OCDE demuestran hasta qué punto la calidad de la enseñanza depende directamente de los docentes y de los centros, cualesquiera que sea el papel de las medidas externas adoptadas por las autoridades educativas. De la Orden (1993: 264) señala a este respecto:

"*La calidad de la educación, en cuanto se manifiesta en un producto válido, dependerá fundamentalmente de lo que acontece en la escuela, de las estructuras y procesos de las instituciones educativas*".

3. EVALUACIÓN Y GESTIÓN DE LA CALIDAD EDUCATIVA

Decíamos al principio del capítulo que la gestión de la calidad educativa era una exigencia actual que había sido producto de cómo había ido cambiando la dinámica de las estructuras formativas y educativas.

En el momento actual, el concepto de evaluación ha sido sustituido por el de Gestión de la Calidad Educativa; como señala Mateo (2000, p.10) "*la nueva cultura de la evaluación ya no se orienta tanto hacia la sanción, la clasificación y la selección como hacía aportar una información razonada y razonable dirigida a la orientación de la gestión de la mejora educativa*".

Este cambio de concepción ha surgido porque el contexto político, económico, social, cultural y científico en el que se tienen que desenvolver las organizacio-

nes educativas ha sufrido un cambio vertiginoso. Vivimos en una sociedad cambiante y en ella la formación y la educación se convierten en elementos estratégicos, de ahí que la mejora de la calidad educativa se convierta en un objetivo fundamental de todos los países desarrollados. Coincidimos con López Rupérez (2000: 45) en que "*si no se da un cambio profundo en el seno de las instituciones educativas en tanto que organizaciones, en su concepción de la gestión y de las prácticas correspondientes, la educación española tendrá serias dificultades para adaptarse a los nuevos tiempos, lograr mejores cuotas de calidad para todos y contribuir sustancialmente al progreso personal, social y económico*".

En este marco social, es necesario tener un sistema de referencia con cierta capacidad explicativa desde la cual se pueda pensar, diseñar e implementar las propuestas relativas a la gestión de la calidad educativa asociadas a la mejora escolar. Para ello, vamos a exponer los cuatro pilares fundamentales que dan sentido a este nuevo paradigma organizativo: la concepción epistemológica, dimensión ética, dimensión orientada a los resultados y propuesta metodológica.

El gráfico que presentamos a continuación, basándonos en las aportaciones de Rupérez (2000) y que nosotros reinterpretamos, sintetiza los distintos referentes que vamos a aportar:

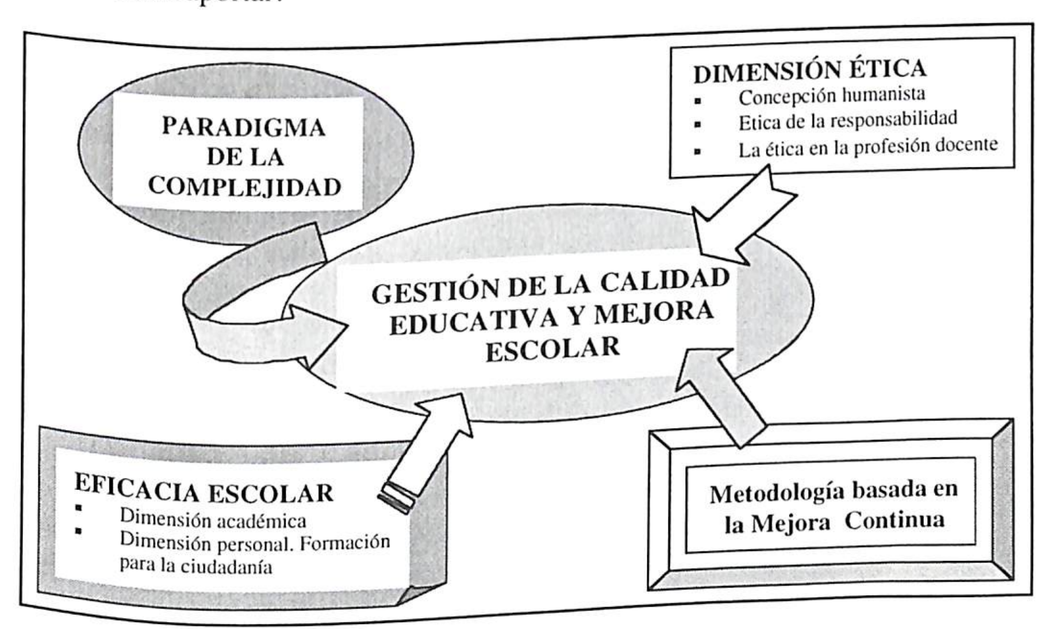


Gráfico 1. Gestión de la Calidad Educativa y Mejora Escolar.

- ***Marco epistemológico:*** la concepción epistemológica que subyace en el movimiento a la gestión de la calidad educativa, asume los planteamientos básicos del *enfoque sistémico*. Las características fundamentales del enfoque sistémico se articulan en torno a: a) un cambio en las relaciones entre las partes y el todo. La

perspectiva analítica según la cual el todo se constituye por la adición de las partes, es superada por un planteamiento que asume las posibilidades de la sinergia, es decir, el refuerzo de diferentes acciones causales individualmente consideradas para dar lugar a un efecto cuantitativamente mayor del que cabría esperar por la suma de los efectos parciales, b) papel preponderante que se otorga a las relaciones causales de carácter circular y en general a las interrelaciones a la hora de comprender la realidad, c) la articulación de la realidad en niveles de complejidad creciente. Cada nivel posee sus propiedades específicas y un comportamiento propio, lo que no obsta para distinguir relaciones verticales, o entre niveles, que tienen su repercusión en las relaciones horizontales, o internas a un nivel dado y d) la revalorización de lo contextual y lo relacional. Por una parte, las relaciones que el sistema establece con su entorno y que condicionan su comportamiento y evolución y por otra la conceptualización de la realidad en términos entre entidades que son afectadas por otras en el seno de los sistemas complejos organizados. En definitiva, este enfoque sistemático en el panorama de la ciencia contemporánea se fundamenta en el *paradigma de la complejidad* (Morin, 1991). Desde este enfoque paradigmático, la percepción que se tiene de los sistemas educativos es como sistemas formados por una gran cantidad de agentes independientes, de distintos tipos, interaccionando entre sí de múltiples maneras, a diferentes niveles organizativos y de acuerdo con leyes más o menos complicadas; de acuerdo a esta percepción, surge la convicción de que la reflexión sobre el sistema educativo, sobre sus reformas y, en general, sobre las políticas que le son características, no pueden ignorar esa nueva lógica, esa nueva forma de pensar la realidad física, natural y social a lo largo de las dos últimas décadas. Para Lévy-Leblond (1991) los sistemas complejos conjugan una *heterogeneidad estructural*, en el sentido de que se manifiesta en ellos la existencia de una estructura jerarquizada en niveles de organización, con una *reciprocidad funcional* que alude a la existencia de acciones recíprocas entre elementos y entre niveles, a circuitos de retroalimentación, y, en general, a conexiones de carácter circular. La aproximación que acabamos de describir es característica del movimiento por la gestión de la calidad educativa que se distingue por su enfoque global, integral o sistémico y asume tácitamente los rasgos anteriormente descritos. En el plano metodológico, para ser organizaciones que aprenden nuestra labor investigadora se orienta hacia la creación de contextos donde determinados mecanismos se manifiestan con cierta regularidad; esta regularidad entendida a modo de cuasi-ley constituye el marco del conocimiento científico que se genera en las organizaciones educativas en la actualidad (Pawson y Tilley, 1997).

• ***Dimensión ética:*** las organizaciones educativas de calidad se caracterizan por un entramado coherente de valores individuales y de valores compartidos que marca el estilo organizativo, orienta las decisiones, da fuerza a los compromisos y otorga sentido a las actuaciones de todo el personal. Los valores centrales que han de acompañar a las organizaciones de calidad se agrupan en racimos éticos.

Entre ellos señalamos citando a López Rupérez (2000) los siguientes:

a) Concepción humanista de las relaciones tanto externas como internas de la organización.
b) Revalorización de la ética de la responsabilidad.
c) Actualización de la ética de la profesión docente.

Desde una concepción humanista, en las instituciones de calidad, las relaciones con las personas se fundamentan en los principios de respeto a la dignidad del individuo, de lealtad, es decir, de corrección ética y de confianza recíprocas. Se proponen estimular la faceta de compromiso y lo consiguen depositando confianza en las personas y creando un clima de reconocimiento, de motivación y de refuerzo, tanto de la dirección hacia los profesores como de éstos hacia los alumnos, mediante la transferencia de expectativas positivas, la valoración de sus logros, la aceptación de sus ideas y la exaltación de los buenos resultados. Además esta concepción humanista ha de alcanzar asimismo, a la esfera externa de relación y particularmente a las familias y a la sociedad en general. Desde esta concepción la noción mercantilista de *cliente* se ha humanizado para poner el acento en la atención a las necesidades humanas de los beneficiarios de un bien o de los usuarios de un servicio con criterios de calidad. La mejora de las relaciones externas de una organización educativa constituye una exigencia ética; su credibilidad y legitimidad social no es algo que esté asegurado para siempre sino que requiere de una renovación permanente que permita garantizar un elevado grado de satisfacción del entorno social en el que desarrolla su actividad (Cortina, 1997; García Maza, 1997). Para generar este entramado de relaciones externas e internas es necesario una revalorización de la *ética de la responsabilidad* que tome en consideración no sólo los principios o el valor intrínseco de las actuaciones sino también los resultados, las consecuencias previsibles de la propia acción. Esta revalorización atañe a la institución educativa en su conjunto pero también afecta a los planteamientos individuales e introduce una modulación en cuanto a una ética renovada de la profesión docente de acuerdo a una deontología profesional actualizada. La incorporación activa a un proyecto compartido constituye, en el contexto actual, una nueva exigencia ética de la profesión docente.

• ***La importancia de los resultados de la educación:*** la idea de la eficacia escolar se asienta en el panorama educativo internacional. Las iniciativas desarrolladas por las políticas educativas denotan que en la actualidad la educación tiene que asumir el papel de explicar el origen de las diferencias en el rendimiento de los alumnos y de las instituciones educativas en general. Esta idea que nace en los años setenta con el movimiento de escuelas eficaces cobra ahora un sentido nuevo: las organizaciones educativas de calidad han de ser organizaciones eficaces, es decir, aquellas que alcanzan los fines, las metas y los objetivos que la sociedad espera de ellas. La noción de eficacia escolar en este contexto se abre y no se constriñe únicamente a los resultados académicos sino que se extiende hacia la formación de nuestros alumnos

para el ejercicio de una ciudadanía responsable y de los valores que constituyen el eje fundamental de la ética de la responsabilidad. Desde este nuevo enfoque de eficacia escolar la equidad cobra también un nuevo significado: "Sólo una escuela será equitativa si es eficaz, es decir, si por su buen funcionamiento contribuye a compensar las diferencias debidas a factores socioeconómicos o socioculturales".

- ***Dimensión metodológica:*** a nivel metodológico la gestión de la calidad educativa se inspira en el concepto de mejora continua; es lo que se conoce como ciclo de Deming o ciclo PDCA (Plan-Do-Check-Act). Consta de cuatro fases fundamentales:
  - — En la *fase de planificar* se identifica el problema y se definen sus características en el marco de una actuación preliminar de diagnóstico; a partir de lo cual se elabora un plan de actuación guiado por ciertas hipótesis basadas en hechos o datos.
  - — En la *fase de hacer* se desarrolla lo planificado poniendo en marcha ordenadamente las acciones previstas en el plan.
  - — En la *fase de verificar* se confronta lo obtenido con lo planificado y se efectúa una evaluación tanto de los procesos como de los resultados, a fin de completar una imagen causal fiable de los aciertos y de los errores.
  - — En la *fase de actuar* se implementan las actuaciones necesarias para la corrección de errores y la consolidación de los aciertos. Todo el conocimiento teórico y empírico generado a lo largo del desarrollo del ciclo se pone a disposición para la ejecución del ciclo siguiente.

El gráfico que mostramos a continuación muestra el ciclo de la calidad de acuerdo a la propuesta metodológica de Deming.

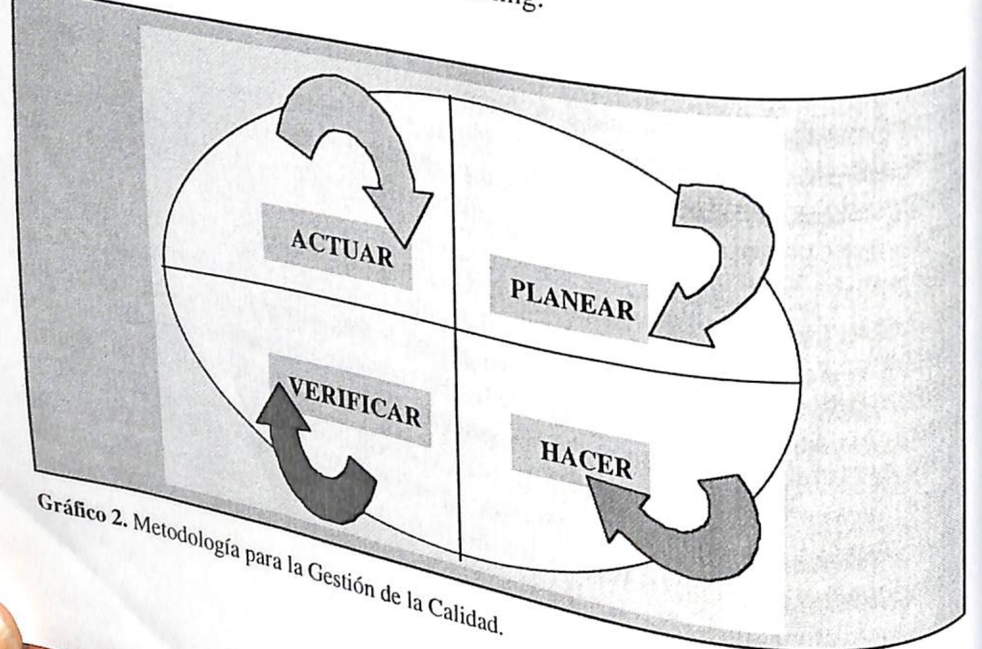


Gráfico 2. Metodología para la Gestión de la Calidad.

Si bien la implantación de esta metodología constituye un objetivo terminal en toda organización educativa como muestra el gráfico nº 3, aún estamos lejos en nuestro contexto de que exista de manera generalizada; es por ello por lo que a partir de ahora nos vamos a centrar en el concepto y herramientas de la gestión de la calidad como heurística que hasta ahora se ha revelado útil para modular las relaciones entre las organizaciones y la sociedad en general en un plano de mejora.

3.1. Gestión de la Calidad Educativa

La gestión de la calidad educativa es una estrategia organizativa y una metodología de gestión que hace participar a todos los miembros de una organización con el objetivo fundamental de mejorar continuamente su eficacia, eficiencia y funcionalidad.

El gráfico que mostramos a continuación tomado de Sallis (1993) ilustra cómo ha ido evolucionando en el tiempo este concepto de calidad en relación con las instituciones educativas:

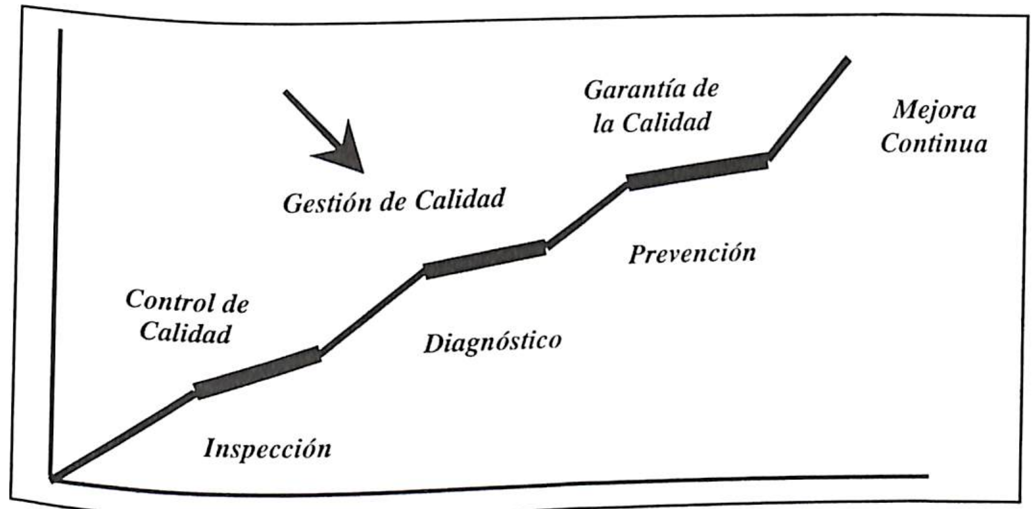


Gráfico 3. Evolución del concepto de calidad en relación a las instituciones educativas.

En sus comienzos, si hacemos un somero repaso histórico al tema del control de calidad, las estrategias de aseguramiento de la calidad se centraban en inspeccionar los resultados, los productos. Se denominaba "control de la calidad del producto". En las etapas intermedias, se centraban en comprobar la uniformidad de los standards de proceso y servicio para toda la producción. Eran los sistemas de aseguramiento de la calidad de la empresa (control de procesos); la revolución de la calidad apareció a principios de los años 50 cuando Deming empezó a centrarse en el control de la calidad durante el proceso de producción del producto y no tras él.

Es lo que se denominó como "construir la calidad". Este mismo concepto trasladado al contexto empresarial estadounidense abogó por integrar en manos del director, el control de calidad, la planificación de la calidad y la mejora de la calidad. Fue en Estados Unidos cuando se empezó a poner el acento en la gestión de la calidad como un concepto más amplio del cual el control estadístico era sólo una de sus herramientas. Esta noción queda superada por el Control de la Calidad Total que implica a todos los elementos de la organización.

3.1.1. *Principios y características fundamentales de la Gestión de la Calidad Educativa*

La implantación de un sistema de calidad en una organización educativa se fundamenta en los siguientes principios que son compartidos por la mayoría de los autores aun cuando existen diferencias en cuanto a la importancia relativa de cada uno de ellos.

a) Proceso orientado a la satisfacción de las necesidades y expectativas de los principales destinatarios.
b) La mejora permanente de todo lo que la organización pretende alcanzar. Para que se desarrolle esta actitud es necesario explicitar los objetivos de la organización y la visión de lo que ésta desea alcanzar.
c) La garantía de la calidad de los procesos internos. Esto implica que se crean normas, se definen los procedimientos para alcanzarlas y se garantiza el cumplimiento de los mismos. La garantía eficaz de la calidad contiene en sí el germen de la mejora permanente; al contrario, una garantía de calidad ineficaz conduce a la burocracia.
d) Orientación al proceso. Una vez determinados todos los procesos esenciales para la organización, la calidad final de un producto o servicio depende de todos los procesos precedentes.
e) Prevenir en lugar de supervisar para lograr la calidad. A partir de unas medidas preventivas adecuadas (planificación correcta, formación selectiva, comunicación eficaz, utillaje adecuado, etc.) se producirán menos errores y se necesitará un menor grado de inspección y control.

A éstos añadimos una serie de principios operativos que son los que marcan la diferencia en aquellas organizaciones que se inspiran en los principios de la gestión de la calidad:

a) *Compromiso y dirección del equipo directivo de una organización.* El equipo directivo ejerce una fuerza impulsora que debe transmitirse al resto de los demás miembros y a todos los niveles de gestión dentro de la organización.
b) Es necesario *realizar un trabajo en equipo eficaz e intencional*. Se convierte dentro de un sistema de calidad tanto en una necesidad como en un resultado de las estrategias eficaces que despliega el sistema.

c) *La calidad es tarea de todos*. La calidad no es una responsabilidad funcional o específica de un departamento, es necesaria la participación de todos los miembros en todos los niveles y departamentos. Debe apoyársele mediante estrategias de recursos humanos que reconozcan y recompensen los esfuerzos por la calidad.
d) *Centrarse en los hechos*. En este aspecto es importante el desarrollo dentro de la organización de un pensamiento factual a partir de la aplicación de procedimientos diversos para recoger información válida y fiable en el que poder fundamentar la toma de decisiones.
e) *Supone la resolución sistemática de problemas*. En este contexto el concepto "problema" debe interpretarse como todo aquello que puede mejorarse; para ello es necesario utilizar procedimientos y herramientas adecuadas que permiten identificar los puntos débiles de la organización.

Este marco operativo necesita un contexto de aplicación. Para ello, en primer lugar es necesario declarar cuál es la *visión y misión de la organización* de una manera consensuada; ésta determinará la política y estrategia de calidad a seguir. En segundo lugar, la creación de un *manual de calidad* que describa la organización, su política, procesos esenciales y la responsabilidad y autoridad de todos los implicados en la organización. Finalmente este proceso tiene que ir acompañado de la *formación sistemática de todos los miembros* de la organización. La formación tiene que ir acompañada de políticas de recursos humanos que recompensen el perfeccionamiento, el desarrollo profesional y los avances en la calidad.

3.1.2. *Aportaciones de la Gestión de la Calidad a contextos formativos*

La mejora en sí misma no es el único motivo que puede barajar una organización educativa para decirse a implantar un sistema de calidad. En general se plantean cuatro argumentos fundamentales que pueden servir de orientación para tomar la decisión (Van der Berghe, 1998: 27):

a) promocionar una imagen de calidad y credibilidad de la institución
b) dar respuesta a factores externos
c) desarrollar un sistema completo de aseguramiento de la calidad
d) la mejora de actividades específicas de la organización.

Cualquier tipo de organizaciones educativa y formativa necesita ofrecer una *imagen de calidad*. Esta imagen se consolida si existe un compromiso de la institución con la calidad; para mostrar una imagen de calidad es necesario a su vez empezar a desplegar iniciativas que muestren al exterior el compromiso efectuado. Este flujo entre compromisos y acciones dependientes de los compromisos se proyecta al exterior; si además este proceso está validado por un agente externo independiente la sociedad en general se volverá aún más exigente con las ofertas formativas y el proceso siempre deviene en las demandas efectuadas a la institución.

A veces la espiral de la calidad descrita anteriormente viene propiciada por una *presión externa* (demanda específica de una organización, expectativas implícitas de otras organizaciones, organismos gubernamentales, un requisito de las organizaciones para impartir formación, etc.) pero si no es así es probable que se convierta en el factor que impulse en los próximos años ese proceso. La preocupación social por la calidad de la educación es cada vez mayor.

Para que una organización educativa pueda desarrollar un *sistema complejo de aseguramiento de la calidad* desde los principios expuestos anteriormente se debe centrar en:

- Considerar que el cliente es lo primero.
- Basar la gestión en la mejora continua, en la constante evolución de los procesos.
- Comprometerse y participar, cooperando en lugar de compitiendo.
- Ser conscientes de que el papel del líder es básico.
- Convertirse en una organización inteligente, que aprende.
- Basar la toma de decisiones en datos objetivos.
- Centrarse en pequeños cambios.
- Establecer un sistema de recompensas que premie el esfuerzo.

Además una organización de calidad debe:

- Optimizar el tiempo dedicado en los procesos de formación.
- Prevenir en lugar de detectar fallos.
- No tener miedo de fallar.
- Contemplar las necesidades de los diferentes destinatarios y del mercado.
- Buscar el desarrollo compartido entre proveedores y clientes.
- Reconocer la responsabilidad social de nuestras acciones.
- Invertir gran cantidad de tiempo.
- Ser ágiles al transmitir la información.
- Generar una cultura institucional compartida.
- Emprender programas de formación.

El que una organización educativa se haga eco de este proceso supone realizar mejoras específicas de la calidad (un mejor control de los formadores externos, desarrollo de un procedimiento minucioso para el diseño, desarrollo y evaluación de los programas formativos, etc.).

3.2. Herramientas para la Gestión de la Calidad: Normas de Calidad

La implantación de un sistema de calidad en una organización educativa requiere de unas herramientas que posibiliten el desarrollo del sistema. Las herramientas fundamentales de un sistema de calidad son las normas de calidad. Antes de su desarrollo, vamos a realizar una distinción conceptual que delimite las dife-

rencias fundamentales que existen entre la garantía de la calidad, las normas de la calidad y la gestión de la calidad total.

El término *garantía de la calidad o aseguramiento de la calidad* se define como todas aquellas actividades planificadas y sistemáticas realizadas dentro del sistema de calidad y que demuestran ser necesarias para asegurar correctamente que una organización cumplirá los requisitos de la calidad; si bien la garantía de calidad es sólo uno de los principios básicos de la gestión de la calidad su consecución se halla estrechamente vinculado a otros elementos de la misma (orientación al cliente, la mejora continua, orientación hacia los procesos y actitud preventiva frente a los problemas que se planteen).

No es el objetivo de este apartado hacer un análisis exhaustivo de las normas de calidad pero si vamos a trazar ciertos perfiles sobre todo por las posibilidades reales de aplicación al ámbito formativo.

3.2.1. *Las normas ISO 9000 de garantía de la calidad*

Las normas de garantía de calidad "ISO 9000" es el nombre común para un grupo de normas internacionales para la gestión de la calidad dentro de las organizaciones. Su origen se remonta al desarrollo de la industria militar americana a principios de los años 50; el aumento de la capacidad productiva de ese país, la exigencia de mayores requisitos de seguridad así como el empléo a gran escala de proveedores, hicieron necesario crear "normas de calidad militar". Este proceso se hizo extensivo a toda Europa y en 1987 la Organización Internacional de las Normas (International Standars Organization-ISO) publicó un grupo de normas denominadas "ISO 9000" de tal manera que este conjunto de normas se convirtieron en normas europeas para la garantía de la calidad. En 1994 se realizó una primera revisión de las mismas y actualmente está en proceso otra que atiende a la estructura, una orientación más hacia la gestión de la calidad que hacia la garantía de calidad y con una terminología aplicable a todos los sectores económicos (en la página web de la Asociación española para la normalización y certificación -AENOR-, se puede consultar las diferencias fundamentales con respecto a la vigente). La revisión que se está realizando quedará de la siguiente forma: ISO 9000: Sistemas de gestión de Calidad. Conceptos y Vocabulario, ISO 9001: Sistemas de Gestión de la Calidad: Requisitos y ISO 9004: Sistemas de Gestión de la Calidad: Guía para llevar a cabo la mejora. Las actuales normas ISO 9001, ISO 9002 y ISO 9003 se integrarán en una única norma ISO 9001.

Concebidas en un principio para utilizarse únicamente en el sector productivo y en las relaciones contractuales proveedor-cliente, con el paso del tiempo, distintos tipos de organizaciones empezaron a interesarse por su uso y de facto se ha convertido en la norma industrial de garantía de calidad para cualquier tipo de organización; si bien su uso en el campo de la enseñanza y la formación es todavía

restringido, probablemente será un proceso ineludible para todas las organizaciones educativas en un espacio corto de tiempo. Las razones que dimos al inicio de este capítulo así como las políticas de calidad que se han puesto en marcha en nuestro país en distintos ámbitos de gestión nos hacen pensar en este aserto. Este grupo de normas pueden considerarse como especificaciones generales de aseguramiento de la calidad, con elementos adicionales tomados de la gestión de la calidad total.

Las normas no tienen un carácter prescriptivo, sino que exigen que una organización: a) defina y planifique sus procesos, b) los documente de manera correcta, c) compruebe su actitud, y d) garantice el control y revisión de los mismos.

Los principios sobre los que se articula este conjunto de normas son los siguientes:

a) la organización posee objetivos de calidad claros
b) existen acuerdos claros entre todos los participantes
c) la organización posee los recursos necesarios para conseguir el nivel requerido de calidad
d) la organización define por sí misma los procesos y recursos necesarios para la calidad
e) todos los procesos y sistemas están sometidos a controles, con evaluaciones y modificaciones cuando sea conveniente
f) todo lo necesario para garantizar la calidad se halla documentado
g) los registros de la calidad permiten una verificación y una comprobación de la garantía de la calidad.

En la página siguiente vamos a presentar una tabla que aglutina el contenido de las normas ISO en sus dos versiones: la de 1994 y la que se está realizando actualmente que sin ser definitiva realiza una aportación fundamental con respecto a la anterior: el concepto de mejora continua con el objetivo de responder mejor a las necesidades y expectativas de los principales usuarios o clientes.

La interpretación que de estos términos se realiza aplicada a contextos educativos y/o formativos queda recogida en la tabla nº 4. La traslación realizada no es la única posible por ello queremos plantearlo a modo de propuesta y sugerencia para las organizaciones educativas.

Los trabajos realizados en el contexto europeo por el Centro Europeo para el Desarrollo de la Formación Profesional (Cedefop) muestran como la aplicación de estas normas de calidad se hacen más plausibles en unas organizaciones educativas que en otras. En relación con la enseñanza obligatoria o superior el sector de la formación profesional y la formación continua resultan ser niveles educativos en los que resultan más fácil su implantación por estar mas sujetos a las leyes de la oferta y demanda laboral. La experiencia desarrollada en el País Vasco en los últimos cinco años en los centros de formación profesional es también una muestra de su mayor factibilidad (1999).

NORMAS ISO 9000 (VERSIÓN 1994)	NORMAS ISO 9000 (VERSIÓN 2000)
• **Principios generales de un sistema de calidad** – Sistema de calidad. Generalidades – Política de calidad – Representante de la dirección – Recursos – Responsabilidad y autoridad • **Requisitos generales para la implantación del sistema de la calidad** – Procedimientos del sistema de calidad – Planificación de la calidad – Formación – Control de los documentos y los datos – Control de los registros de la calidad • **Principios generales del control de procesos** – Control de los procesos – Inspección y comprobaciones – Técnicas estadísticas – Control de productos suministrados por los clientes – Identificación y trazabilidad de productos – Estatus de la inspección y las pruebas • **Requisitos específicos para procesos y actividades particulares** – Revisión de contratos – Control del diseño – Compras – Control de los equipos de inspección, medición y comprobaciones – Manipulación, almacenamiento, embalaje, conservación – y entrega – Servicios postventa • **Mantenimiento y mejora del sistema de calidad** – Control de los productos no conformes – Intervenciones correctoras y preventivas – Control interno de la calidad – Revisión por la dirección	Se han articulado en torno a cuatro capítulos básicos. Éstos son: • **Responsabilidad de la dirección** • **Gestión de los recursos** • **Realización del producto** • **Medición, análisis y mejora** Además se han introducido otros capítulos de carácter general que indican: • **Alta dirección** Se ha puesto un mayor énfasis en el papel de la alta dirección, lo cual incluye su compromiso en el desarrollo y mejora del sistema de gestión de la calidad, consideración de los requisitos legales y reglamentarios, y el establecimiento de objetivos medibles en todas las funciones y niveles relevantes de la organización. • **Mejora continua** Se requiere que las organizaciones dispongan de un proceso de mejora continua, dentro de su Sistema de Gestión de la Calidad • **Satisfacción del Cliente** Se requiere que una organización tenga un enfoque al cliente y haga un seguimiento de la satisfacción del mismo. Este seguimiento de la satisfacción es un medio que debe utilizarse para evaluar las prestaciones del Sistema de Gestión de la Calidad. • **Comunicación interna** Se requiere que una organización disponga de un proceso de comunicación interna, para suministrar información sobre el Sistema de Gestión de la Calidad y su efectividad • **Interacción entre procesos** Se requiere que la organización defina sus procesos y como interaccionan entre ellos • **Recursos** Los requisitos incluyen la necesidad de evaluar la eficacia de la formación, suministro de la información relevante, comunicación interna y externa, necesidad de instalaciones y los factores físicos y humanos del entorno de trabajo que puedan afectar a la calidad del producto • **Exclusiones permitidas** Se ha incluido el concepto de "exclusiones permitidas" como una vía para tratar con el amplio espectro de organizaciones y actividades • **Terminología** Se han realizado cambios en la terminología para reflejar el vocabulario utilizado por las organizaciones. El término "organización" sustituye al término "suministrador" que se utilizó previamente para referirse a la unidad a la cual era de aplicación la Norma Internacional.

Tabla 3. Contenidos fundamentales de las normas ISO 9000 en las dos versiones actuales.

Términos de las ISO 9000	Interpretación para una Organización Educativa o Formativa
• **Proveedor**	Diseñador o impartidos de la enseñanza o la formación
• **Clientes**	Estudiantes/alumnos en el caso de la enseñanza; empresarios o formados en el caso de la formación
• **Producto**	Curso, programa educativo o formativo; material educativo o formativo
• **Cuadros ejecutivos**	Director general/ director del centro de enseñanza/rector/comité de directores/...
• **Contrato**	Todos los tipos de acuerdos con los clientes, incluso los implícitos, desde la matriculación de los estudiantes para un curso, hasta los acuerdos negociados de formación específica para un cliente.
• **Diseño**	Definición de especificaciones educativas/formativas; diseño de programas de cursos y currículos; especificación de los contenidos del material educativo/formativo; diseño de instrumentos de evaluación para estudiantes/ alumnos/formados.
• **Compras**	Adquisición de recursos materiales y humanos y servicios necesarios, incluyendo el uso de profesores/formadores/adjuntos temporales, independientes o externos
• **Procesos**	Se refieren al desarrollo, la planificación y la impartición de enseñanza o formación, incluyendo en ella todos los elementos fundamentales del currículum formativo.
• **Inspección y Pruebas**	La evaluación de cursos/programas formativos/materiales educativos por los estudiantes, alumnos formados y empresarios que los contratan
• **Calibración**	Se refiere a la validez de los procedimientos de evaluación utilizados
• **No conformidades**	Todo problema que surja durante el desarrollo y la impartición de enseñanza o formación
• **Formación**	Formación y desarrollo profesional de los miembros de la organización educativa o formativa

Tabla 4. Trasposición de los conceptos de las normas ISO 9000 al ámbito educativo.

4. LA GESTIÓN DE LA CALIDAD EDUCATIVA EN ESPAÑA: POLÍTICAS DE CALIDAD

Desde que empezó en nuestro país este interés creciente por la evaluación, la calidad y de manera consecuente por la gestión de la calidad educativa las instancias políticas han puesto en marcha iniciativas y planes estratégicos que se han ido traduciendo de manera paulatina en acciones más o menos directas que a su vez han ido configurando políticas de calidad en distintos ámbitos de aplicación.

No es nuestro objetivo hacer una presentación exhaustiva de todas ellas por ser la temática central de la segunda parte de esta publicación pero sí queremos presentar cómo las concepciones y planteamientos que se han ido mostrando a lo largo de este capítulo han tenido una proyección en la política educativa de nuestro país a la vez que muestran el valor institucional de las decisiones adoptadas.

4.1. Políticas de calidad en los centros educativos

En relación a los centros educativos, la política iniciada por el Ministerio de Educación y Cultura en relación a la gestión de la calidad de los centros educativos se ha desarrollado a partir de la implantación en todo el territorio Mec del Modelo Europeo de Gestión de Calidad (EFQM).

La cronología de la política de evaluación de centros en nuestro país ha seguido el proceso que recogemos en la tabla siguiente:

MARCO LEGISLATIVO	FECHAS	PLAN DE EVALUACIÓN	INSTANCIA QUE EVALÚA
Art. 62 Título IV de la Logse	1991-92	Proyecto Piloto del Plan Eva (Plan de Evaluación de Centros)	Inspección Central del Ministerio de Educación y Ciencia y el CIDE
	1992-96	Aplicación extensiva del Plan EVA	
Art. 29 Título III Ley Orgánica Lopegce	1998 (Boe, nº 131 de 2 de Junio)	Modelo Europeo de Gestión de Calidad en los centros docentes	Ministerio de Educación y Cultura. Dirección General de Centros Educativos
	1998 (Boe, nº 141, de 13 de Junio)	Orden del Plan Anual de Mejora de los centros docentes	Ministerio de Educación y Cultura. Dirección General de Centros Educativos

Tabla 5. Cronología de la política de Evaluación de Centros en España.

Desde el Ministerio de Educación y Cultura se puso en marcha en 1996 un conjunto de políticas tendentes a promover la mejora de la calidad en los centros educativos. Se trataba de introducir en el ámbito de la educación pública los principios de la gestión de la calidad y la mejora continua. Este conjunto de políticas se estructuró en tres niveles de complejidad organizacional creciente: la administración educativa, los centros educativos y el aula. De forma paralela se puso en mar-

cha como eje transversal una política de formación dirigida fundamentalmente a inspectores, administradores de la educación, directivos y candidatos a directores escolares.

En las Comunidades Autónomas con competencia en educación la política de evaluación de centros se ha articulado a partir de la creación de un departamento o agencia de evaluación que realiza las siguientes funciones: a) la evaluación de los centros escolares no universitarios, evaluación del rendimiento escolar y la evaluación del rendimiento del sistema educativo, con funciones parecidas a las que realiza el Instituto Nacional de Evaluación y Calidad. Los aspectos concretos de cómo se desarrollan los planes de evaluación de centros quedan recogidos en el capítulo 5 de esta publicación.

El resultado de esta política de carácter evolutivo queda expuesta en la tabla nº 6 que contempla los distintos ejes sobre los que se extendió esta política de calidad.

4.2. Política de calidad en la Institución Universitaria

La política de calidad en la institución universitaria se desarrolla través del Plan Nacional de Evaluación de la Calidad aprobado el 25 de septiembre de 1995 por el pleno del Consejo de Universidades. El Real Decreto 1947/1995 de 1 de Diciembre (B.O.E. 9-12-95) eleva a texto normativo dicha propuesta.

La tabla nº 7 ilustra cómo ha ido evolucionando en nuestro país la política de evaluación en el sistema universitario.

Las características más destacables del plan desde el punto de vista de la política de la evaluación son a juicio de Rodríguez Espinar (1998: 148) las siguientes:

- La propuesta evaluativa que se realiza supera la controversia entre rendición de cuentas versus mejora de la calidad; la finalidad de la evaluación desde las orientaciones que propone el plan es hacia la mejora de la calidad.
- Un concepto evaluativo que se desarrolla a partir de la contextualización de los procesos evaluativos y del desarrollo en ese contexto de una cultura evaluativa.
- La concesión de ayudas económicas para llevar a cabo el proceso de evaluación acorde con los proyectos presentados por las propias universidades. La primera convocatoria se realizó en 1996 y hubo una respuesta masiva de las distintas universidades.
- Crear un marco de referencia único dentro de todo el territorio español consensuado con todas las comunidades autónomas que tienen competencia en educación. El objetivo es la consolidación de un modelo de evaluación del sistema universitario español.

	Plan Anual de Mejora	Modelo EFQM	Formación en Gestión de la Calidad	La Gestión de la Calidad en la Administración Educativa periférica
Curso 1996-97	• Se pone en marcha el Plan Anual de Mejora con la participación de 268 Centros. • Se publican las instrucciones correspondientes en el BOMEC.	• Se adopta el Modelo Europeo a las características de los Centros públicos. Se elaboran los materiales correspondientes.	• Se incorpora la Gestión de la Calidad a la formación de Directivos e Inspectores. • Se desarrollan las I Jornadas sobre Enseñanza Pública y Gestión de la Calidad.	• Se plantea la conveniencia de introducir la Gestión de la Calidad en las unidades que se relacionan directamente con los centros.
Curso 1997-98	• Se regula la segunda edición mediante Resolución de la Dirección General de Centros Educativos. • Se difunde ampliamente el Plan. • Se duplica el número de ayudas a los mejores planes.	• Se publican los materiales correspondientes. • Se publican a todos los centros del territorio MEC. • Se aplica de forma experimental el modelo en dieciséis centros de seis direcciones Provinciales.	• Se incorpora la perspectiva organizacional a la formación para la acreditación. • Se extiende la formación en Gestión de la Calidad. • Se desarrollan las II Jornadas sobre Enseñanza Pública y Gestión de la Calidad.	• Se inician en algunas Direcciones Provinciales experiencias de Gestión de la Calidad en sus unidades administrativas.
Curso 1998-99	• Se consolida la regulación sobre Planes Anuales de Mejora mediante orden Ministerial.	• Se triplica el número de centros que adoptan el Modelo Europeo. • Se generaliza su estructura de Autoevaluación como referencia para el diseño por los centros de su evaluación interna. • Se ofrece a las Comunidades Autónomas el conocimiento generado para su eventual utilización. • Se convoca el 1er Premio Nacional a la Calidad en Educación.	• Se consolida la Gestión de la Calidad como elemento de formación de directivos, inspectores y administradores de la educación. • Se organiza un Congreso Europeo sobre Enseñanza Pública y Gestión de la Calidad.	• Se ofrece a las Comunidades Autónomas la posibilidad de avanzar en el proceso de implementación de una Gestión de Calidad en este tipo de unidades.

Tabla 6. Política de Gestión de la Calidad de Centros Educativos impulsada por el Ministerio de Educación y Cultura (López Rupérez, 2000: 61).

	SISTEMA POLÍTICO	SISTEMA UNIVERSITARIO	
		Actividades Consejo de Universidades	Actividades Unión Europea
1978-1985	Constitución Ley de Reforma Universitaria Creación del Consejo de Universidades		
1986-1989			
1990-1991		Estudios previos para definir un standard	
1992-1993		Debate sobre el alcance de la evaluación orientados a mejorar la calidad	
1994-1995		Programa Experimental de Evaluación	
1996-1997		Valoración del Programa Experimental. Acuerdo del Consejo de Universidades por el que se propone un plan de Evaluación Coordinación del Proyecto Piloto Europeo	Proyecto Piloto Europeo Conferencia final del Proyecto Europeo
1998	1ª Convocatoria del Plan Nacional 1ª Convocatoria de Comunidades Autónomas	Coordinación del Plan de Evaluación 1er Informe de Evaluación Propuesta de 2ª Convocatoria	Propuesta de recomendación de la Comisión... de garantía de calidad en la enseñanza superior Dictamen del Consejo Económico y Social. Modelo Europeo de Evaluación

Tabla 7. Cronología de la política de evaluación (De Luxan, 1998).

4.3. Políticas de calidad en el sistema educativo

Desde su creación en 1993 (R.D. 928/1993) del Instituto Nacional de Evaluación y Calidad (INCE) en el marco normativo de la Logse la política de calidad del sistema educativo se ha centrado en la aprobación de tres planes de actuación que desde 1994 han iniciado proyectos de carácter nacional e internacional con el objetivo fundamental de velar por la calidad de nuestro sistema educativo y realizar estudios comparativos entre los países de la Unión Europea y con Iberoamérica para establecer pautas de acción que respondan a una globalización de los sistemas educativos.

De una forma sintética los planes de actuación aprobados dentro de la estructura organizativa del INCE han posibilitado desarrollar los proyectos expuestos en la tabla de la página siguiente.

Estos proyectos culminan con la realización de informes evaluativos que permiten pulsar a los gobiernos, agentes educativos y a la sociedad en general el funcionamiento del sistema educativo desde una perspectiva multidimensional.

Proyectos	Plan de Actuación 94-97	Plan de actuación 98-99	Plan de actuación 99-00
Acciones de carácter permanente	• Elaboración de un sistema estatal de indicadores de la calidad de la educación	• Elaboración de un sistema estatal de indicadores de la calidad de la educación • Construcción de un banco de ítems y pruebas de rendimiento • Diseño y desarrollo de programas de formación de especialistas en evaluación	• Elaboración de un sistema estatal de indicadores de la calidad de la educación • Construcción de un banco de ítems y pruebas de rendimiento • Diseño y desarrollo de programas de formación de especialistas en evaluación • Edición de la Revista de Educación
Proyectos específicos	• Evaluación de la Educación Primaria • Evaluación Piloto de la Educación Secundaria Obligatoria • Evaluación Piloto de la Formación Profesional	• Estudio sobre la enseñanza y aprendizaje de lenguas extranjeras en España • Evaluación de la Educación Primaria (2º estudio) • Actuaciones para promover la evaluación y la calidad en la gestión de los centros públicos	• Estudio sobre la enseñanza y aprendizaje de la lengua inglesa en España • Evaluación de la Educación Primaria (2º estudio) • Evaluación de la Educación Secundaria Obligatoria (4ª curso) • Diagnóstico de la función directiva en los centros docentes públicos
Estudios Internacionales	• Tercer estudio internacional de Matemáticas y Ciencias (TIMSS) • Las Lenguas en la Educación (LES) • Estudio Hispano-Francés sobre la enseñanza y el aprendizaje de las lenguas extranjeras	• Proyecto INES y Proyecto PISA de la OCDE • Proyecto piloto en la Unión europea sobre evaluación de la calidad de la educación • Evaluación de la calidad de la educación en la cumbre Iberoamericana en el seno de la OEI	• Proyecto INES y Proyecto PISA de la OCDE y estudios de centros de Educación secundaria post-obligatoria • Evaluación de la calidad de la educación en la cumbre Iberoamericana en el seno de la OEI

Tabla nº 8. Planes de actuación del Ince desde su creación en 1993.

ACTIVIDADES PARA EL ALUMNO

1. Analiza dentro de un sistema de calidad las distintas acepciones que puede adoptar "el producto educativo" y plantea otras posibles:

> a) Producto= los servicios educativos o formativos impartidos, incluyendo instrumentos y servicios relacionados.
> b) Producto= adquisición de aprendizaje.
> c) Producto= el estudiante, formado o alumno.
> d) Producto= el título o cualificación obtenidos.

2. Diseña un sistema de aseguramiento de la calidad en el contexto universitario de acuerdo al ciclo PDCA.
3. Analiza la interpretación de los términos de las normas ISO 9000 para una organización educativa que se ha realizado en el texto (tabla nº4). Plantea otras interpretaciones posibles.
4. Una vez estudiadas las normas de calidad realiza un cuadro sinóptico que recoja las principales ventajas e inconvenientes que representa la implantación de un sistema de aseguramiento de la calidad en una organización formativa.

BIBLIOGRAFÍA COMENTADA

CANO GARCÍA E. (1998). *Evaluación de la Calidad Educativa.* La Muralla: Madrid.

Este texto aporta una síntesis sobre distintos ámbitos de la calidad en educación. Comienza con un recorrido histórico sobre los estudios de calidad, analiza el concepto de calidad y su tratamiento en los documentos de la administración educativa. Posteriormente se diversifica en los distintos ámbitos de la calidad. Su orientación es sobre evaluación de la calidad más que sobre gestión de la calidad, pero para iniciarse en el tema resulta una referencia clara y sincrética.

CLUB GESTIÓN DE CALIDAD (1998). *Relación existente entre aseguramiento de calidad (ISO 9000) y gestión de calidad total.* Edita Club Gestión de la Calidad: Madrid.

Esta publicación del Club de gestión de la Calidad aporta un análisis comparativo entre las normas de calidad ISO 9000 y el Modelo Europeo de Gestión de la Calidad (EFQM), que permite comprender la interrelación existente entre ambos y más ampliamente clarifica la creciente confusión sobre lo que representa Aseguramiento de la Calidad y Gestión de la Calidad Total.

PÉREZ JUSTE, R., LÓPEZ RUPÉREZ, F., PERALTA, M.D. Y MUNICIO, P. (2000). *Hacia una Educación de Calidad. Gestión, Instrumentos y Evaluación.* Madrid: Narcea.

Este texto de reciente aparición combina la aportación de distintos profesionales del campo de la educación y con distintas aproximaciones teóricas y metodológicas. Se estructura en cuatro grandes capítulos: el concepto de calidad educativa, la gestión de la calidad, los instrumentos de la calidad y la evaluación de la calidad. Es fruto de un seminario permanente celebrado durante dos cursos académicos sobre calidad en educación y organizado por la fundación Castroverde.

REY GARCÍA, A. (1998). *Cómo gestionar la Calidad en las Universidades.* Edita Club Gestión de la Calidad: Madrid.

Este texto es una aplicación práctica del Modelo Europeo de Excelencia Empresarial en la Universidad pública, que evidencia la forma de impulsar iniciativas innovadoras en una amplia gama de actividades y la utilidad de los modelos empresariales y mejora de la institución universitaria.

VAN DE BERGHE, W. (1997). *Aplicación de las Normas ISO 9000 a la enseñanza y la formación. Interpretación y orientaciones desde una perspectiva europea.* Luxemburgo: CEDEFOP.

Este texto es un documento elaborado por el Centro Europeo para el Desarrollo de la Formación Profesional sobre la calidad en la Formación Profesional. El trabajo no se limita a presentar las normas ISO 9000 sino que también interpreta cada apartado de la misma desde una perspectiva europea. Contiene asimismo sugerencias prácticas para su aplicación y resalta los requisitos necesarios antes de implantar un sistema de calidad. Su aportación fundamental reside en la adaptación que se ha realizado de estas normas aplicadas a contextos formativos.

FOROS NACIONALES E INTERNACIONALES DE DISCUSIÓN Y DIRECCIONES ÚTILES EN RELACIÓN CON LA GESTIÓN DE LA CALIDAD EDUCATIVA

ORGANISMO	INFORMACIÓN PROPORCIONADA
Club Gestión de Calidad http://clubcalidad.es	Organización sin ánimo de lucro fundada en 1991 por empresas españolas líderes en su sector de negocio y comprometidas con la Gestión de la Calidad Total.
Club Gestión de Calidad en Andalucía http://www.iat.es	Esta organización tiene una proyección en Andalucía mediante un acuerdo con el Instituto Andaluz de Tecnología.
Asociación Española para la Calidad http://www.aec.es	Esta web ofrece información sobre su centro de información, la revista Calidad, así como datos sobre Congresos y exposiciones.
Aenor. Centro de Formación http://www.aenor.es	Asociación Española de Normalización y Certificación. Organismo Nacional de Normalización. Primera entidad acreditada por ENAC para la certificación de sistemas ISO 9000 e ISO 14001, verificación medioambiental y QS 9000.
Infocalidad http://infocalidad.net	Web dedicada a los profesionales del mundo de la calidad, con multitud de información y base de datos.
Consejeros en Gestión de Personal. (CGP) http://www.cgpnet.com	Una extensa guía de recursos sobre calidad; entidades de normalización y certificación.
Hispacal http:/www.hispaclub.com/hispacal	Es un proyecto desarrollado sin ánimo de lucro gracias a la colaboración desinteresada y solidaria de profesionales de la calidad del ámbito iberoamericano. El objetivo fundamental de Hispacal es mejorar el conocimiento sobre la Calidad en el ámbito Iberoamericano.
Sociedad Latinoamericana para la Calidad (SLC) http://www.calidad.org	Foro Latinoamericano para la Calidad, que integra en su web varios servicios, como fondos bibliográficos, eventos, alojamientos, etc.
Centro Europeo para el Desarrollo de la Formación Profesional (Cedefop) http://www.cedefop.gr	Es un instituto europeo que ayuda a los responsables políticos y los profesionales de la práctica formativa, ya pertenezcan a la Comisión Europea, a los Estados Miembros o a los agentes sociales de toda Europa, a fundamentar sus decisiones sobre política de formación profesional. Puede proporcionarle las informaciones más recientes sobre la situación actual y las tendencias futuras en la formación profesional de la Unión Europea.

SEGUNDA PARTE

LA GESTIÓN DE LA CALIDAD EDUCATIVA EN LA EVALUACIÓN DE PROGRAMAS, CENTROS E INSTITUCIONES Y SISTEMAS EDUCATIVOS



CAPÍTULO III

CALIDAD Y EVALUACIÓN DE PROGRAMAS: USOS Y DISEÑO DE LA EVALUACIÓN

José María Gobantes Ollero[1]
Universidad de la Laguna

OBJETIVOS

Los objetivos de este capítulo son los siguientes:

1. Identificar los distintos significados del concepto de calidad y su sentido cambiante según pretensiones e intereses acerca de los programas.
2. Relacionar las diferentes concepciones del concepto de calidad con las decisiones a tomar acerca de las finalidades y el modelo de evaluación a desarrollar.
3. Mostrar la vinculación entre el creciente interés por la calidad y la revisión a la que está siendo sometida la finalidad del sistema educativo y los programas sociales.
4. Definir el concepto de programa y mostrar la importancia de que sea la evaluación la que se adapte a la complejidad del mismo.
5. Exponer la idea de que la evaluación no debe ser menos compleja que el programa evaluado ni debe someterlo a presiones que desvirtúen sus propias finalidades.
6. Identificar los principales cambios, y sus causas, producidos en la teoría y en la práctica de la evaluación de programas.
7. Aportar los elementos necesarios para que el lector se sitúe entre las ventajas y las desventajas de los modelos de evaluación resumidos.
8. Animar a realizar el trabajo intelectual de elaborar propuestas creativas que resuelvan las necesidades prácticas dadas en cada circunstancia y contexto local.
9. Proponer los elementos, la estructura y el proceso a seguir para la elaboración sistemática y creativa de un diseño de evaluación de programas.
10. Transmitir la importancia de elaborar criterios de referencia para retomarlos en el análisis riguroso de los datos y evidencias obtenidos. Describir la forma de elaborarlos y la forma de usarlos en el análisis.
11. Mostrar la necesidad de introducir en (y desde) el diseño de evaluación la planificación de las *estrategias de evaluación* para la implicación de las audiencias y el uso práctico de los resultados.
12. Guiar la elaboración de un ejercicio práctico de diseño de evaluación de programas de intervención social con la finalidad de iniciar al lector en dicha tarea profesional.

Actividades para el alumno
Bibliografía comentada

[1] Dpto. Didáctica e Investigación Educativa y del Comportamiento. Área de Métodos de Investigación y Diagnóstico en Educación. Centro Superior de Educación. Universidad de La Laguna. Campus Central. C/ Delgado Barreto s/n. Tfno: 922319218, E-mail: jgobante@ULL.ES

CONTENIDOS

Calidad y Evaluación de Programas: uso y diseño de la evaluación

Calidad y Evaluación
Dependiendo de la concepción de calidad, la evaluación tendrá uno u otros fines, objetivo, ...
Empezar por aclarar operativamente el concepto de calidad. Usos del concepto:
- Prestigio/excelencia
- En función de los recursos
- Resultados/productos
- Valor añadido
- Adecuación a propósitos
- Perfeccionamiento o mérito (De Miguel, 1995).

Concepto de programa
Conjunto de acciones humanas y acciones materiales diseñados e implantados sistemáticamente en una determinada realidad con el propósito de resolver algún problema. Pero tal grado de definición previa no siempre se da y sin embargo no podemos renunciar a que la evaluación se lleve a cabo y a que sea de utilidad en tales condiciones.

3. Evaluación: evolución conceptual. Tres periodos
Evaluación por objetivos y experimental
A partir de la Conferencia de Cambridge en 1972: iluminativa, respondente, autoevaluación democrática, etc.
A partir de las críticas de Stenhouse a la evaluación iluminativa: problemas con los criterios de evaluación, la interpretación y la implicación de las audiencias.

4. Estrategias de evaluación
Incluir en el diseño operaciones que impliquen a las audiencias a lo largo de todo el proceso de evaluación.

5. Elementos y estructura del diseño.

Fase I.
Finalidad y objetivos, papel del evaluador y análisis del programa.

Fase II.
Características de la evaluación.
Autojustificación

Fase III.
Medios y recursos, dimensiones, indicadores, técnicas y métodos, agenda, análisis y discusión, elaboración de informes...; metaevaluación.

6. Actividades para los alumnos

INTRODUCCIÓN

Este capítulo tiene por finalidad introducir al lector en las claves de la evaluación de programas. La aproximación conceptual realizada de manera genérica sobre el término calidad, en el capítulo anterior, cobrará en éste mayor especificidad aplicada a la evaluación de programas. Por ello, se argumenta que la elección de las características técnicas y proceso de la evaluación está condicionada por la concepción (criterios) que se tenga acerca de qué es calidad y qué se demanda como calidad al programa que se pretende evaluar. En este sentido se alerta acerca del poder que la evaluación puede llegar a tener, tanto para mejorar como para desvirtuar las finalidades y resultados del programa que se evalúa.

Por otro lado, además de explicar el concepto de programa de intervención social, proponemos los elementos, la estructura y el proceso a seguir para la elaboración sistemática y creativa de un diseño de evaluación de programas. Hacemos hincapié, por un lado, en la importancia de la sistematización de la planificación, aun sabiendo que el proceso introducirá cambios necesarios en la misma, y, por otro lado,

en la necesidad de planificar desde el diseño, las estrategias que se pretenden realizar para lograr la implicación de las audiencias (en mayor o menor grado dependiendo de la finalidad de la evaluación) y el uso práctico posterior de los resultados. A su vez, para ello, insistimos en la importancia de elaborar marcos de referencia (criterios) compartidos y fundamentados que sean retomados para el análisis de los datos y evidencias. Finalmente se propone el ejercicio de elaborar un plan de evaluación de un programa, elegido previamente como motivo práctico y utilizando como apoyo el contenido aportado por el capítulo. La finalidad del ejercicio que se propone es introducir al lector en la práctica profesional de elaboración de propuestas prácticas de evaluación, cada vez más demandadas socialmente.

1. CALIDAD Y EVALUACIÓN DE PROGRAMAS

Actualmente es ineludible tener que referirnos al concepto de calidad al hablar de educación y, sobre todo, al hacerlo de la evaluación de programas. Aunque no siempre ha sido así. El término *calidad* aparece muy tardíamente de forma explícita, como objetivo relevante del sistema educativo (en 1987), con las propuestas de reforma que posteriormente darían lugar a la LOGSE. Ello no significa que anteriormente no hubiese preocupación acerca de la calidad de la educación. Tampoco ha sido tradicionalmente una preocupación exclusiva de los agentes externos a las instituciones o a los programas, tales como políticos, gestores, administradores, economistas.

La importancia que el concepto tiene para la evaluación (que es el punto de vista que nos interesa aquí) es que, según sea la concepción de calidad de la que se parta, o lo que es lo mismo, según lo que se espere del programa evaluado, a la evaluación se le propondrá unos u otros objetivos, se orientará hacia la búsqueda de unos u otros indicadores y establecerá unos u otros procesos y estrategias como ahora veremos. Es por ello que, para evaluar, es necesario empezar por aclarar lo más ajustadamente posible en qué se está pensando cuando se demanda la evaluación de la calidad.

El uso del concepto de calidad está en auge, y se habla de ella en cualquier contexto o tema del que se trate: servicio de calidad, ventas de calidad, compras de calidad, calidad de vida, calidad de producción, calidad en el diseño, relaciones humanas de calidad... Es una palabra mágica sobre la que todo el mundo parece estar de acuerdo acerca de su significado pero que, sin embargo el mismo término denomina una disparidad de criterios en cuanto a lo que significa en la práctica. En una sociedad caracterizada por la pluralidad y cada vez más fragmentada, es lógico que en la demanda de lo que la evaluación tiene que valorar, nos encontramos con tantas interpretaciones de calidad como concepciones existen sobre la función de la educación y lo que se espera de ella, ya sea desde una perspectiva teórica, social o personal.

Las distintas audiencias (planificadores, gestores, personal que lo lleva a la práctica, destinatarios, etc.) de los programas de intervención social trasladan sus intereses y concepciones acerca de lo que deben ser los programas (educativos o no) a criterios que juzgarán su calidad. Así, la diversidad de referentes hace muy difícil lograr unificar operativamente los criterios acerca de lo que significa calidad. La OCDE (1991: 39) señala esta dificultad del siguiente modo:

> *"El hecho de que la educación se haya visto cada vez más politizada en los últimos años y que a menudo ya no sea la regla un amplio consenso nacional sobre los objetivos, puede ser considerada como la razón de que la preocupación por la calidad en la escolarización se haya situado en el primer plano y como causa principal de que su determinación sea tan problemática".*

Clarificar el papel que se le demandará a la evaluación dependerá del difícil logro del consenso señalado. Será el evaluador el que en cada contexto deberá decidir la finalidad o finalidades de la evaluación, la diversidad de criterios a los que deberá atender y, en consecuencia, los datos e información que necesitará para realizar la evaluación. Al respecto, De Miguel (1995), como vimos en el capítulo anterior, identifica seis usos frecuentes del concepto de calidad (como prestigio/excelencia, en función de los recursos, como resultados/productos, como cambio o valor añadido, como adecuación a propósitos, perfección o mérito). Según señala este autor, el concepto de calidad esta basado en la consistencia de las cosas bien hechas. Para ello es necesario la existencia de una cultura de la calidad y procesos de autoevaluación.

Ante la diversidad de perspectivas, como señala De Miguel (1995: 33),

> *"La falta de una teoría sobre la calidad y las dificultades para poder llegar a un consenso entre las distintas audiencias implicadas impide que se pueda establecer una definición de este concepto que sea aceptada de forma unánime".*

Los distintos puntos de vista, con respecto a qué significa calidad son, socialmente, cada vez mayores y más distantes. Ya sea porque al sistema educativo y a los programas sociales se les pide que atiendan y resuelvan de forma inmediata cada vez más problemas sociales (violencia, drogadicción, paro, maltrato, sexismo, embarazos en la adolescencia, formación para el empleo y hasta el caos del tráfico), cuya profundidad y buena parte de sus raíces están fuera del alcance de la institución escolar y de los programas sociales; sea porque se considera que el esfuerzo de inversión realizado no está dando los resultados esperados; o sea por otras muchas causas posibles, la realidad con la que nos encontramos es que se está haciendo una

revisión en profundidad de la institución escolar y los programas sociales, acerca de la finalidad que la modernidad les confirió como instrumento de culturización, compensación y vertebración social. Como hemos señalado anteriormente, una sociedad cada vez más plural y abierta, que incluye multitud de perspectivas, puntos de vista, ideologías y proyectos de vida diferentes, pero que también es cada vez más desestructurada, parece haber perdido la capacidad para unificar y compartir fines y criterios, que sean realmente educativos, hacia los que dirigir los esfuerzos de mejora. Si la educación no tiene más objetivos que el posibilitar salidas en el mercado de trabajo (la educación reducida a entrenamiento), cualquier escuela es válida si cumple unos requisitos de calidad en esa dirección y si se adecúa a la relación costo-beneficio. Como señala Gimeno Sacristán refiriéndose a la escuela pública,

> *"¿Si no hay horizontes claros, si no hay visión social que defender ni sociedad que vertebrar con proyectos comunes a través de la escuela pública, qué más da un sistema que otro salvo sus diferencias en calidad que pueden perjudicar a la escuela pública?"* (Gimeno Sacristán, 1996: 63).

Se pone en cuestión la finalidad compensatoria, culturizadora, modernizadora, promocionadora de la educación y los programas sociales, y se buscan nuevas exigencias de calidad externa a la propia lógica de las instituciones y programas evaluados pero eficaces en su regulación y reorientación hacia otras prioridades.

La presión externa sobre los programas, para que éstos respondan a criterios de calidad distintos a los educativos se ejerce a través de resortes burocráticos, técnicos o ambos a la vez, como si a través de ellos fuese posible conformar previamente (burocrática o técnicamente) el espacio de indeterminación que el ejercicio docente siempre lleva implícito y que sólo puede ser cubierto con el desarrollo profesional de los docentes y, en general, de los profesionales que desarrollan en la práctica los programas de intervención social.

A pesar de la diversidad de perspectivas acerca de la calidad, como las que acabamos de ver, podemos distinguir en ellas básicamente dos enfoques contrapuestos que demandarán dos fines y usos distintos de la evaluación. La relación entre concepciones acerca de la calidad y evaluación la establece De Miguel (1995: 39) del siguiente modo:

> *"La relación entre teorías sobre la calidad y criterios para evaluarla es tan estrecha que cabe hablar de tantos modelos de evaluación como enfoques teóricos se formulan al respecto".*

De los seis conceptos de calidad anteriormente señalados, propuestos por De Miguel (1995), podemos agrupar los cinco primeros en lo que podríamos denomi-

nar *criterios de calidad externos a los programas*. La utilidad de su uso es comparar instituciones o programas, o valorar el estado de los sistemas educativos o de servicios sociales. Presentan dificultades para reflejar la complejidad de la realidad y la práctica de los programas e instituciones y son insatisfactorios para las personas que desarrollan los programas en contextos locales con características singulares. Los criterios externos no reflejan la compleja realidad de los programas y someten a éstos a una tensión distinta a la de su lógica práctica; criterios externos tales como la búsqueda obsesiva de productos previamente establecidos o la satisfacción de los usuarios convertidos en clientes, etc. La consecución de lo que denominamos criterios externos, son difíciles de garantizar de forma infalible por las personas que trabajan en las instituciones o programas que son evaluados desde dichos criterios.

Por otro lado, y en sintonía con la última concepción de calidad, de las seis que hemos señalado anteriormente, podemos denominar *criterios de calidad internos* al programa evaluado, a aquellos que se centran en el mérito de las prácticas concretas teniendo en cuenta las diversas características de los contextos locales en los que se aplican creativamente los programas.

Mientras que la evaluación, realizada a partir de criterios externos a la práctica del objeto evaluado, tendrá como finalidad establecer la distancia entre lo pretendido previamente (por la legalidad, por los objetivos, por el producto esperado, por las demandas de los clientes, etc.) y lo finalmente logrado; la evaluación, realizada desde la lógica propia del programa y su puesta en práctica (lo que hemos denominado *criterios internos*) tendrá por finalidad y uso aportar información contrastada a todos los implicados por el programa acerca de su funcionamiento concreto y real, que les faculte para elaborar estrategias concretas de mejora del objeto evaluado y les implique en su puesta en práctica.

La evaluación entendida como instrumento de mejora del objeto evaluado se pone a prueba, ante los profesionales que desarrollan en la práctica los programas, según sea su capacidad para analizar la realidad en toda su complejidad, no sobrevalorando la búsqueda de resultados a costa de los procesos y las condiciones de las que se parten (entorno cultura, organización del centro, nivel de coordinación y acuerdo en las finalidades compartidas, innovación, Cultura de calidad, etc.), reflejando la totalidad de los problemas de desarrollo del programa en vez de simplificarlos, etc.

Como venimos señalando, el concepto de calidad del que se parta en la situación dada, condicionará el proceso de evaluación, desde los criterios hasta la información a recoger, el análisis y el uso práctico de las conclusiones a las que se llegue. Por ello, establecer de forma operativa qué se entiende por calidad en cada caso, es una tarea que debe incluirse entre las operaciones a realizar por la evaluación. De Miguel (1995: 34) señala al respecto:

> *"En todo proceso de análisis y valoración sobre los fenómenos educativos se debe partir de la definición operativa que en cada caso concreto se establece sobre la calidad de la educación. De ahí que para evaluar un aspecto concreto del sistema educativo debamos partir de los criterios específicos de calidad que se establecen al respecto".*

Ello incluye tanto procesos de trabajo para desarrollar y unificar criterios (con las distintas audiencias o considerar su diversidad) como revisar la teoría y la normativa existente al respecto a partir de las cuales especificar los criterios de calidad para cada aspecto a evaluar y su contexto. Estas tres son las fuentes desde las que elaborar los criterios como marco interpretativo para realizar la evaluación.

2. CONCEPTO DE PROGRAMA

Cuando nos referimos a evaluación de programas, el concepto incluye a todo objeto susceptible de ser evaluado distinto a personas (evaluación de alumnos y profesores), materiales curriculares o instituciones (evaluación institucional y de centros). Cada una de estas evaluaciones tiene sus especificidades y por lo tanto procedimientos distintos. Sin embargo, dependiendo de las características y complejidad del programa a evaluar, la evaluación puede incluir a personas, instituciones y materiales.

La definición de programa que propone Fernández-Ballesteros (1995) centra perfectamente la idea acerca del significado que expresa tal término:

> *"Conjunto especificado de acciones humanas y recursos materiales diseñados e implantados organizativamente en una determinada realidad social, con el propósito de resolver algún problema que atañe a un conjunto de personas"* (p. 24).

Por lo tanto, los requisitos mínimos que el objeto de evaluación tiene que tener para que lo consideremos como tal es que se trate de una acción intencionada, que persigue unas determinadas finalidades e interviene en un contexto social específico para procurar la promoción educativa, sanitaria o social de las personas a las que va dirigido. Por otro lado, el concepto de programa incluye no sólo al documento escrito que lo contiene sino a su desarrollo tal y como se da en la práctica.

Sin embargo, el concepto de programa aquí expuesto es manifiestamente insuficiente si partimos de la consideración de que un programa debe estar perfecto, sistemática y racionalmente prediseñado, debe contar con recursos humanos y materiales en cantidad y calidad ajustada a las metas que se propone, debe llevarse a la práctica siguiendo su planificación previa y debe contar con unos objetivos precisos, realistas, que acierten a expresar sistemáticamente los logros del progra-

ma antes de que éste se aplique a contextos reales, de forma que podamos valorar a partir de ellos si se han alcanzado las metas perseguidas. Lo ideal es encontrarnos con un programa que cuente con un buen diseño, a partir del que podamos proceder a planificar la evaluación de forma igualmente sistemática. Pero tal grado de definición previa no siempre se da y sin embargo no podemos renunciar a que la evaluación se lleve a cabo y a que sea de utilidad en tales condiciones.

Es frecuente encontrarnos, por diversas causas, que los objetivos son poco menos que la expresión imprecisa y general de las vagas intenciones del programa; que, porque en su momento el programa se hizo rutinario entre el resto de operaciones regulares de la institución, se carece de un programa escrito; que es enorme la distancia entre lo previsto en la planificación escrita (proyecto) y lo que realmente sucede en la práctica; o que, simplemente, buena parte de las personas que aplican el programa no conocen el programa sobre el que se supone que están trabajando (Stufflebeam; 1987, p. 178). Creemos que precisamente la evaluación puede contribuir a mejorar todo ello si comienza por reconocer que la realidad es así de compleja, contradictoria y, en cierto modo, imprevisible. Para una buena práctica de la evaluación (su diseño, desarrollo y uso) más que establecer qué es un programa y qué no lo es, quizás sea importante considerar seriamente las tres siguientes cuestiones acerca de los programas a evaluar.

La primera cuestión que hay que considerar seriamente, referida al programa, para el posterior diseño de la evaluación, es el *grado de complejidad* del programa. Complejidad en cuanto a la extensión geográfica e institucional, en cuanto a la diversidad y cantidad de acciones y actividades incluidas en el programa y en cuanto al tópico del que se ocupe el programa: atención a las drogodependencias, innovación educativa, educación afectivo-sexual, formación para la creación de empleo, educación para el consumo, erradicación de la marginalidad social, etc. Con respecto al tópico de que se trate, su evaluación requiere conocerlo lo más a fondo posible, estudiarlo y asesorarse, en su caso, por expertos acerca del mismo. Queremos decir con ello que la actividad evaluadora es distinta según el campo de conocimiento que esté involucrado y para valorar actividades, estrategias, relaciones, coordinaciones entre profesionales, su formación, los resultados, etc. habrá que atender a las peculiaridades y diferentes culturas profesionales de los agentes intervinientes, características del problema que se trata de paliar o del conocimiento y actitudes que se pretenden transmitir, etc.

En cuanto a la extensión del programa, es evidente que no es lo mismo diseñar la evaluación de un programa circunscrito a un centro escolar, en el que no intervienen más sectores profesionales que el profesorado y va dirigido a los alumnos del centro, que diseñar y desarrollar la evaluación de un programa que abarca varios municipios, concejalías, diversas consejerías y sectores profesionales (salud, enseñanza, asuntos sociales, etc.), asociaciones de padres y madres (AMPAS), agrupaciones juveniles, centros, medios de comunicación, etc. Como veremos más

adelante, todo ello determinará el proceso de evaluación en todas sus fases y, sobre todo, será mucho más complejo unificar criterios para establecer la finalidad de la evaluación, el uso que el programa (sus distintas audiencias) hará de ella y, lo que es más importante, será más complejo lograr que, en la práctica la evaluación satisfaga la diversidad de expectativas creadas hacia ella.

En segundo lugar, desde nuestro punto de vista (nos interesan los problemas prácticos de la evaluación y su uso), es importante tener en cuenta que los programas de intervención social, igual que la educación en general y el sistema educativo en particular, *son elaboraciones sociales que responden a necesidades, intereses y finalidades diversas, determinadas socialmente*. En ese sentido el programa a evaluar tendrá la impronta del contexto (político, económico, administrativo o social) en el que se desarrolla y de las personas involucradas en él de una u otra forma (como responsables o receptores) y lo que ellas deseen del programa, esperen, hagan, entiendan o les afecte. Por ello, cada programa se desarrolla en la práctica concreta, bajo formas y peculiaridades diversas puesto que, como hemos señalado, dependen de los contextos y de las personas que en él participan. Así, por ejemplo, un programa de animación a la lectura, desarrollado en un barrio determinado, acabará teniendo las características que las personas involucradas en él le den por su grado de participación, interés, medios con los que cuenta, formación de los monitores que lo desarrollan, instalaciones, apoyo de las instituciones y sus responsables, etc. A su vez, cada sector social involucrado (padres y madres, jóvenes) demandará del programa respuestas diferentes y tendrá criterios diferentes acerca de su calidad. Por su parte, los jóvenes puede que demanden lecturas de interés para pasar un buen rato, los profesores o monitores desearán contar con una biblioteca bien dotada de modo que los usuarios compensen sus deficiencias culturales, habrá padres que simplemente les satisfaga el programa porque se mantiene a los chicos fuera de las calles y al responsable del programa quizás le preocupe sobre todo la gestión económica y su rendimiento, etc.

Llegar a comprender el programa para valorarlo, incluye atender desde el principio a estos aspectos de su complejidad, de manera que evitemos la abstracción simplificadora del mismo, que a través de la evaluación puede ser transmitida como una idea caricaturizada e irreconocible, con la que nadie se sentirá identificado. Todo ello significa para la evaluación que, además de atender las distintas visiones (criterios de referencia) existentes acerca del programa (aspecto que trataremos más adelante), deberá rastrear el significado que le otorga cada uno de los ámbitos en los que se decide, de una u otra forma, el modo en el que finalmente se desarrollará: contextos, ámbitos de decisión política y administrativa, profesionales que lo diseñan y desarrollan, destinatarios de las acciones, etc. Se trata de un análisis que va de lo macro a lo micro, desde la decisión política de ponerlo en marcha, hasta las acciones y actividades diseñadas para los usuarios. Esta cuestión influirá determinantemente en las dimensiones que propongamos para la evaluación y sobre las que más adelante trataremos.

Por último, y como conclusión de los anteriores puntos, es importante tener en cuenta, al diseñar y desarrollar evaluaciones que lo fundamental, cualitativa y cuantitativamente, es el programa evaluado. No entendemos la evaluación como un fin en sí mismo, sino como un instrumento al servicio de la mejora del programa (Aguilar y Ander-Egg; 1992, p.22). La evaluación es un instrumento poderoso que, de gestionarse mal, puede distorsionar fácilmente la finalidad del programa evaluado y su cometido originario. Esta distorsión ocurre bajo distintas formas de expresión ejercidas por parte de la evaluación: interpretando equívocamente los objetivos que el programa pretende, demandando resultados tangibles que el programa no puede garantizar, alienando desde el exterior la interpretación de quienes lo llevan a la práctica y de sus usuarios, cerrando excesivamente criterios de evaluación que acaban condicionando negativamente la práctica, tomando decisiones que deben ser tomadas desde la lógica del programa, etc.

El éxito de la evaluación va a depender, además del acierto de las decisiones técnicas que se tomen para llevarla a cabo, de la capacidad del propio programa (sus responsables y destinatarios) para incluirla entre sus rutinas, interactuando con ella a lo largo de todo el proceso, demandándole información, datos, evidencias y análisis para tomar decisiones, abriéndole un espacio en la estructura organizativa del programa para que se pueda desarrollar, dedicándole tiempo a la evaluación, etc. Sin embargo todo ello ocurrirá si la evaluación demuestra en cada fase y en cada decisión su utilidad como apoyo para la mejora del programa, su sintonía con las finalidades perseguidas y con las características particulares de cada práctica, contexto y actividad. La utilidad de la evaluación se pone a prueba en cada paso y a lo largo de todo el proceso.

Ello quiere decir que la evaluación por sí sola no hace bueno un programa sino que es un instrumento de mejora siempre que el programa lo utilice como tal. En este sentido, en buena parte, la evaluación depende del programa evaluado.

A partir de estas consideraciones acerca de los programas de intervención social proponemos los elementos y los pasos necesarios para elaborar diseños de evaluación que logren responder a las necesidades de los programas y contribuyan a su mejora.

3. EVOLUCIÓN DEL CONCEPTO DE EVALUACIÓN: MODELOS

Queremos prevenir, desde el principio acerca del riesgo de simplificación que implica reducir el trabajo de investigación sobre evaluación y de prácticas muy diferentes realizadas a lo largo de un periodo de tiempo muy amplio, a esquemas que necesariamente tienen que ser lo más sencillos posibles cuando se trata de hacer una primera aproximación a las propuestas de evaluación y sus diferentes modelos.

Prácticamente cada autor aporta enfoques, prácticas y soluciones propias y distintas, por lo que no deberemos perder la riqueza de las diferencias de cada

experiencia y enfoque a la vez que analicemos las semejanzas y diferencias de enfoques acerca de los aspectos más relevantes como, por ejemplo, el concepto de calidad, que acabamos de ver, etc. Como hemos señalado ya, el evaluador en cada situación deberá decidir creativamente cómo sintetizar la experiencia de los distintos modelos de evaluación, de todos ellos, con el propósito de ajustarse eficazmente a la finalidad que la evaluación tenga en cada caso.

Este apartado vamos a dedicarlo, en primer lugar, a trazar la evolución histórica del concepto de evaluación (por la cantidad de textos disponibles que tratan el tema ampliamente no será necesario detenernos particularmente en ello) para, en segundo lugar, proponer dos puntos de corte que establecerán tres enfoques distintos de evaluación. Tres enfoques que determinarán diferencias en aspectos básicos para el diseño y desarrollo de la evaluación como su finalidad y usos, el papel del evaluador o evaluadores, los procesos y estrategias de evaluación, la relación entre evaluadores y evaluados o el producto de la evaluación. Estos cortes serán los criterios que tendré en cuenta para diferenciar unos u otros modelos o enfoques de evaluación y que discutiré ampliamente.

En cuanto a la evolución histórica de la evaluación de programas es comúnmente aceptado establecer las siguientes épocas. La primera 40-50, la segunda en los años 60 y la tercera a partir de finales de los 60 y principios de los 70.

Un primer periodo de referencia es el de los años 40 y 50 en el que destaca la propuesta de evaluación educativa de R. Tyler. Es conocida su definición de evaluación como "el proceso de determinar en qué medida el currículo y la enseñanza satisfacen realmente los objetivos de la educación" (1977: 109). Influido por las teorías conductistas sobre el aprendizaje de la época, la evaluación debía establecer la distancia existente entre los objetivos de los programas y los resultados obtenidos tras la aplicación del programa, a partir de los objetivos previamente establecidos, referidos a los alumnos y explicitados en forma de conductas directamente observables. Las ventajas de la propuesta de Tyler y su gran influencia posterior, perceptible hasta la actualidad, estriban en la superación de las insuficiencias de las anteriores prácticas parciales de evaluación en educación basada en tests y pruebas de la época o costosos diseños experimentales difíciles de centrarse en el propio programa evaluado, en sus finalidades y en sus logros (Ginsburg, 1989). La evaluación a través de los objetivos, los programas eran evaluados directamente a partir de sus propias pretensiones y no de forma indirecta.

Un segundo período de enorme *expansión de la práctica evaluativa y puesta en crisis de las prácticas de evaluación tradicional*. Período situado en torno a la década de los años sesenta y caracterizado, en Norte América, por la preocupación acerca de los dudosos resultados educativos de la gran expansión de las inversiones en el sistema educativo y en programas sociales. Se diseñan programas curriculares a gran escala y, paralelamente, se financia la realización de evaluaciones que verifiquen los resultados obtenidos y justifiquen las inversiones (*accountability*). La ex-

pansión de las prácticas de evaluación de programas de esta época hacen que se revele las insuficiencias técnicas en la detección de la información relevante acerca de lo que ocurría con los programas evaluados y que, por lo tanto no aportaban valoraciones para mejorar la calidad educativa de los mismos.

En esa época, un comité de expertos diagnostica la gravedad de la situación en la que según ellos está sumida la evaluación (Stufflebeam y Shinkfield, 1987). Situación que se caracteriza por procesos de evaluación ambiguos, superficiales, inadecuados y tendenciosos, falta de acuerdos entre los expertos acerca de lo que significaba evaluar y mala preparación de los mismos; ausencia de una teoría pertinente, falta de instrumentos y planificaciones apropiadas e inexistencia de sistemas de organización, procesamiento y presentación de la información de evaluación.

Como producto de la situación descrita y la búsqueda de soluciones, se diversifican enormemente las perspectivas acerca de la evaluación. Mientras por un lado se profundiza en la búsqueda de diseños experimentales que respondan a las necesidades de los programas (Cronbach, 1963) y se elaboran pruebas y tests centrados en la especificidad de cada programa (basados en criterios), por otro lado se recomienda el desarrollo de modelos que distingan a la evaluación de actividades como la investigación experimental, la valoración, el juicio profesional o el análisis acerca de la coincidencia entre objetivos y resultados (Stufflebeam y Shinkfield, 1987: 182).

Por último podemos considerar un tercer período desde principio de los años 70, caracterizado por la profusión de modelos diferentes de evaluación tratando de superar las deficiencias de lo que Tejedor (1993) denomina evaluación tradicional. Entre esos modelos de evaluación destacan el modelo formativo-sumativo y el de la evaluación libre de metas de Scriven (1969 y 1972), el modelo de discrepancia entre los estándares del programa y su realización, de Provus (1971) el modelo CIPP de Stufflebeam (Stufflebeam y Shinkfield, 1987), la evaluación iluminativa de Parlett y Hamilton (1983), la evaluación *de réplica* de Stake (1983 y 1987), la evaluación democrática de MacDonald (1983) o la autoevaluación de H. Simons (1985 y 1995), la evaluación como adaptación del procedimiento judicial (B. Wolf). Estos modelos amplían la perspectiva de la teoría y la práctica de evaluación hasta entonces centrada en la objetividad de los enfoques basados en los objetivos y los diseños experimentales. Los nuevos modelos consideran los criterios y los juicios como rasgos de la actividad de evaluación (Medrano, 1995).

La fecha de referencia es 1972, año en el que se celebra una conferencia en Cambridge a la que asisten, entre otros, Stake, Parlett, MacDonald y Hamilton. Según las conclusiones de la Conferencia, en el futuro la evaluación deberá responder a las necesidades y puntos de vista de las distintas audiencias; clarificar los complejos procesos prácticos de la organización, la enseñanza y el aprendizaje; servir a las decisiones que han de tomar el público y los profesionales.

Por otro lado, proponían, entre otras, que la evaluación debía utilizar la información de la observación cuidadosamente verificada; el diseño de la evaluación

debía realizarse de forma que pudiera adaptarse a los acontecimientos no previstos en la planificación del programa evaluado; la perspectiva del evaluador y sus criterios deberían ser manifestados a las audiencias; el papel del evaluador se debía ampliar al asumir tareas complementarias a las del estrictamente técnico; las decisiones educativas en las que está implicada la evaluación, tiene componentes sociales, políticos y económicos que no se deberían ignorar; los evaluadores debían prestar atención al problema de quién debe hacer la interpretación de los resultados de la evaluación: o el experto evaluador o las diferentes audiencias implicadas.

Como resumen del giro que sufre la teoría de evaluación, ante las deficiencias prácticas que presenta la evaluación basada en los objetivos de los programas, reproducimos en la figura de la página siguiente (Tabla I) una comparación, elaborada por Stufflebeam y Shinkfield (1987) de las características de lo que Stake denomina evaluación preordenada y la evaluación de réplica del propio Stake (1983).

Por lo tanto, durante los años 70 y 80 se produce una tendencia de alejamiento con respecto a la epistemología positivista tradicional hacia paradigmas *naturalistas* (Eisner, 1979; Guba y Lincon, 1981; Lincon y Guba, 1986; Stake, 1978; Willis, 1978). Otra tendencia que se produce en esta época es la sustitución de la preocupación prioritaria por la búsqueda de información, por parte de expertos (el evaluador como experto que establece la verdad de lo ocurrido con el programa por medio de sus métodos expertos experimentales o interpretativos) por la preocupación prioritaria por la comunicación y la implicación de las audiencias en el proceso de evaluación y la incorporación de sus prioridades como prioridades también de la evaluación. En palabras de Weis (1989): "la pregunta ha cambiado de ¿qué conocimiento? a ¿conocimiento de quienes?". La revitalización de la investigación-acción (un ejemplo de evaluación como investigación-acción es el trabajo de Midkiff y Burke, 1987) y la distinción entre evaluación hecha desde fuera de los programas o desde su interior ha hecho que tome fuerza la preocupación por incrementar la utilización de los hallazgos de la evaluación y, en ese sentido, la implicación de los beneficiarios del programa aumenta la consiguiente validez de los hallazgos de la evaluación (Messick, 1989; House, 1993; Brandon, Lindberg y Wang, 1993).

Otra línea de trabajo, para evitar la tendenciosidad en las evaluaciones realizadas por un sólo evaluador es la de las evaluaciones *políticas o de debate* de Owens, Wolf, y Weis (García Ramos 1992 y Medrano et. al., 1995), en ellas diferentes evaluadores, o grupos de evaluadores, ofrecen sus distintos puntos de vista sobre los diferentes aspectos evaluados de un programa. Son los *modelos de perspectiva contrapuesta* o los *modelos de orientación judicial*.

La preocupación, que acabamos de enunciar, acerca de los problemas de la comunicación de la evaluación y la implicación en ella de las personas relacionadas con el programa hace que se cuestionen, cada vez más insistentemente las estrategias de evaluación. La ineficacia del uso de los informes de evaluación al comuni-

Distinción	Evaluación preordenada	Evaluación de réplica
1. Propósito	Determinar hasta qué punto han sido alcanzadas las metas	Ayudar a los clientes a distinguir las virtudes y los defectos
2. Alcance de los servicios	Satisfacer los requisitos informativos tal como se había acordado al principio del estudio	Responder a los requisitos informativos de la audiencia a lo largo de todo el estudio
3. Contratos	Las obligaciones de las partes formales de la evaluación se negocian y definen tan específicamente como sea posible al principio del estudio	Los propósitos y procedimientos se esbozan de manera muy general al principio y se desarrollan durante el estudio
4. Orientación principal	Propósitos del programa, indicación de las variedades	Problemas y actividades del programa
5. Planificación	Preespecificadas	Se van haciendo a sí mismas
6. Metodología	Reflejo del modelo investigativo: intervención y observación	Reflejo de lo que la gente hace naturalmente: observa, interpreta, especifica
7. Técnicas preferidas	Diseño experimental, objetivos de comportamiento, hipótesis, muestreo al azar, tests objetivos, estadísticas de resumen e informes investigativos	Estudio de casos concretos, objetivos expresivos, muestreo intencionado, observación, exámenes de programas opuestos, e informes expresivos
8. Comunicaciones entre el evaluador y el cliente	Formales e infrecuentes	Informales y continuas
9. Bases para la interpretación de los valores	Referencia a los objetivos preestablecidos, a los grupos de reglas y a los programas competitivos	Referencias a las distintas perspectivas de valor de la gente que se tiene a mano
10. Diferencias fundamentales	Sacrifica el servicio directo a los que toman parte en el programa por la producción de informes investigativos objetivos	Sacrifica cierto grado de precisión en la valoración para que pueda aumentar la utilidad
11. Previsiones para evitar la tendenciosidad	Utiliza procedimientos objetivos y perspectivas independientes	Reproducción y definición operacional de los términos ambiguos

Tabla I. Principales distinciones entre la evaluación preordenada y la evaluación de réplica de Stake (tomado de Stufflebeam y Shinkfield, 1987: 254).

car los resultados y su aplicación práctica para la mejora lleva a plantear el reparto de papeles y el protagonismo de los evaluadores y las personas evaluadas. Nos detendremos brevemente en ello.

Para mejorar la comunicación y, con ella la implicación, los modelos de evaluación que se desarrollan como alternativa a la evaluación basada en los objetivos de los programas, ponen gran atención al problema de la comunicación entre evaluación y programa, entre evaluadores y evaluados. Ello les lleva a centrar su atención en cuestionar y revisar la utilidad de los informes tradicionales. En esta línea J. Weis (1989) señala como uno de los males de la evaluación el uso de informes tradicionales que no conectan con quienes desarrollan los programas. Según el autor citado, el problema de comunicación para la implicación se reduce a escribir bien. Es decir, para conseguir el uso del informe de evaluación en la práctica hay que mejorarlo incluyendo en el estilo de escribir las técnicas literarias y usar formatos que atraigan a los lectores. Por otro lado, el mismo autor señala que hay que añadir:

> *"aquellas experiencias vividas de un modo que permita a aquellos que están más implicados tener acceso a estas experiencias. Esto es, proporcionar un contexto para que los individuos reflejen sus circunstancias y se cuestionen lo que se da por hecho. Un proceso tal de experiencia cuestionadora puede incluir una comprensión de la forma en que son las cosas, del porqué han llegado a ser de esta manera, de los límites para el cambio y de qué oportunidades pueden darse para hacer que las cosas sean de otro modo".*

En cuanto al mismo problema, Stake (1987) dedica un artículo a la elaboración de informes ante la reconocida insuficiencia de los formatos tradicionales. Stake mantiene la misma línea de argumentación acerca del valor de los informes que el autor anterior, es decir, si no se leen los informes es porque están mal elaborados:

> *"Bien, (¡eso es terrible!). No el hecho de que una persona no haya leído el informe sino el hecho de que quien quiera que hiciese el estudio de evaluación no lo convirtió en valioso para esa persona. La responsabilidad de la evaluación es una responsabilidad de comunicación. (...) Por tanto es la oportunidad y la obligación de cada evaluador de programas traducir los requisitos de evaluación, dados por el gobierno estatal o federal, en un plan que recoja información útil para las personas que lean el informe, particularmente la gente local".* (Stake, 1987: 315).

La importancia conferida a los informes (finales o no) de evaluación es la evidencia de la importancia del problema, no resuelto, de la comunicación de la

evaluación para lograr la implicación de las audiencias en las posibles soluciones sugeridas por la evaluación realizada. Esta relación deficiente entre la evaluación y su objeto es expuesta crudamente por Fiddy y Stronach (1987: cit. por Holly y Hopkins, 1988: 230):

> *"La pregunta entonces es: ¿por qué estamos nosotros (como evaluadores, como investigadores dentro de la práctica educativa) tan poco interesados en las relaciones pedagógicas entre nosotros y nuestros clientes? Especialmente dado que la evidencia parece ser que nuestros clientes aprenden poco de nosotros y nos valoran menos aún".*

Entiendo que el problema es tan importante que es difícil pensar que con un buen informe se pueda resolver. Desde mi punto de vista, tras el problema de la baja utilidad práctica de la evaluación aparecen, como telón de fondo deficiencias en las estrategias que desde estas perspectivas se proponen para realizar la evaluación. Las estrategias de evaluación deben revalorizarse para pasar de la idealización teórica a la acción práctica. Es la crítica que Stenhouse hace a lo que él llama la nueva ola de la evaluación (1984a: 82; 1984b: 164). Stake (1987) define muy bien esta cuestión aunque sin ofrecer solución satisfactoria puesto que eso exigiría, desde nuestro punto de vista, un giro acerca de su concepción de la evaluación:

> *"...permítanme cuestionar el valor de los informes finales. Permítanme ser escéptico por un momento. (...) Si la evaluación es evaluación formativa, es decir, evaluación en el nivel del desarrollo; si el propósito de la evaluación es proporcionar control de calidad, mantener el programa funcionando; si el propósito de la evaluación es servir a los de dentro, a la gente que realiza el programa; entonces nosotros no necesitamos, creo yo, mucho en cuanto a informes formales de evaluación. Pero cuando la audiencia está fuera necesitamos preocuparnos mucho más sobre la forma de nuestro informe formal".* (pág. 314).

Este mismo punto de vista, acerca del informe, es sostenido por Holly y Hopkins (1988: 230) al afirmar rotundamente que,

> *"Los frutos de la evaluación, normalmente son el informe escrito y sin embargo ése es un medio particularmente inapropiado para comunicar acerca del desarrollo. Éste es un problema que los «investigadores aplicados» y «los evaluadores del cambio» han estado enfrentando durante algún tiempo y aunque hemos llegado a ser bastante buenos en cuanto a definir la cuestión, identificar soluciones efectivas continúa siendo idiosincrásico y problemático".*

J. Weis (1989) también llega a la conclusión de que para lograr que la evaluación sirva a los individuos, incluidos en una situación determinada, como instrumento de comprensión y de toma de conciencia de su situación y acción para la transformación,

> *"no es suficiente escribir informes sino más bien producir las condiciones para que aquellos cuyas prácticas están siendo estudiadas puedan comprometerse en un encuentro dialógico con el informe. Willis proporciona esta oportunidad al comprometer a los jóvenes de clase trabajadora en la discusión para animarles a que lleguen a saber la forma en que ellos están implicados en lo que les ocurre y para cuestionarles sobre cómo las cosas podrían ser de otro modo. Este intento emancipador es un ejemplo de lo que Lather (1986a) llama la «validez catalítica»: «el grado hasta el cual el proceso de investigación reorienta, centra y energitiza a los participantes hacia el conocimiento de la realidad para transferirla»".*

La larga cita que sigue resume las críticas que Stenhouse (1984b) hace a los modelos de la llamada nueva evaluación y, por lo tanto, algunos de los cambios que deben imprimirse en su teoría y en su práctica:

> *"Me sigue pareciendo que la nueva ola de evaluadores se preocupa por el mérito o valía de un currículum o una práctica educativa, pero sus criterios no están claros y su preocupación por los receptores y la presentación de resultados me parece que enmascara el problema. Aspiran a decir las cosas tal como son y con frecuencia escriben como si ello fuera posible, en tanto que se permiten alguna distorsión debida a sus propios valores. Pero no hay un decir las cosas como son. Existe sólo una creación de significado mediante el uso de criterios y estructuras conceptuales. La tarea de impartir instrucciones a aquellos que asumen decisiones, en un lenguaje que entiendan fácilmente, puede dar lugar a la importación casual de creencias y criterios no bien examinados. Puede resultar tentador el deseo de lograr respuesta por parte de la audiencia, en especial si ésta es poderosa desde el punto de vista político. Y es muy fácil entonces que la evaluación que aspire a la condición de novela degenere en una novelucha".* (pág. 164).

Anteriormente señalábamos la reunión de Cambridge, en 1972 organizada por Parlett y Hamilton, como punto de ruptura entre la evaluación del viejo estilo (Stufflebeam & Shinkfield; 1987: 327), basada en los diseños experimentales, en el producto y en los objetivos; y la evaluación de estilo iluminativo (Stufflebeam & Shinkfield; 1987: 327).

Ahora podríamos considerar las críticas reseñadas de Stenhouse a la forma de concebir la evaluación de autores como Parlett y Hamilton (1983), Stake (1983),

MacDonald (1983) como el punto de inflexión entre la perspectiva centrada en la distinción entre programa planificado y desarrollado y, por otro lado, la perspectiva interesada en la implicación de los destinatarios y los que desarrollan el programa en la práctica.

Por el significado que, tiene para la teoría de evaluación nos detendremos brevemente en el análisis de la cita de Stenhouse. En ella el autor condensa un buen número de críticas. En primer lugar, como señalan Stufflebeam & Shinkfield (1987), Stenhouse sugiere que la evaluación debe concentrar su atención en la construcción crítica de significados más que en calcular el valor y el mérito de lo evaluado, como siguen haciendo los modelos de evaluación a los que critica. En segundo lugar, Stenhouse señala como un problema para el desarrollo de la evaluación el querer responder (evaluación 'respondente' de Stake) a las preguntas que a las distintas audiencias les interesa, como si la verdad pudiera ser establecida al margen de intereses, dice el autor. Centrar la evaluación en ese cometido revela una concepción epistemológica acerca del conocimiento y la verdad que coloca al evaluador en un lugar indefinido entre, por un lado, como instrumento pasivo y neutral, dedicado a trasvasar información de un círculo de interés a otro y, por otro, como el depositario de la verdad de lo que realmente ocurre y a la que se puede acceder a través del conocimiento de las cosas (expresado por el evaluador de forma literaria y atractiva en el informe final). En tercer lugar, Stenhouse señala la imposibilidad de lograr todo esto por medio del informe de evaluación, aspecto este sobre el que ya nos hemos extendido suficientemente. En cuarto y último lugar, plantea el problema del poder y su problemática relación con la evaluación y su desarrollo.

Como conclusión de este breve repaso de la evolución de los modelos de evaluación[2] de programas (en la tabla II se presenta un resumen de los modelos más importantes de evaluación y sus características más significativas). Hay que señalar al respecto que no se trata de sustituir unos métodos de evaluación por otros. Entendemos que el período de la contraposición entre modelos de evaluación y constructos metodológicos (cuantitativos versus cualitativos), en cierto modo ha sido superada. Actualmente el reto estriba en ajustar los diseños de evaluación a las exigencias de los programas y, así como los programas son cada vez más diversos, extensos y complejos, la evaluación también tiene que reflejar esa complejidad incluyendo diversidad de procedimientos y metodologías. Debe, para ello, comenzar por establecerse la finalidad y el uso (o usos) que desde el programa se le exige. Es muy difícil que un sólo modelo por sí mismo y en estado puro pueda responder

[2] Para la profundización en los diferentes modelos de evaluación de programas el lector puede consultar, además de las obras originales de los autores citados, los siguientes manuales: BRINKERHOFF et al. (1985); WORTEN & SANDERS (1987); STUFFLEBEAM Y SHINKFIELD (1987) y COLÁS REBOLLO (1993).

DESCRIPTORES	***Congruencia entre resultados y objetivos***	***Orientados a la toma de decisiones***	***Orientados al juicio***	***Políticos o de debate***	***Intuición pluralista u holísticos***
Representantes	*Tyler* *Metfessel y Michel* *Popham* *Cambell y Stanley* *Suchman*	*Cronbach* *Stufflebeam* *Alkin*	*Scriven* *Stake* *Eisner*	*Owens* *Wolf, R. L.* *Weis*	Stake Parlett y Hamilton Rippey MacDonald
Organizadores previos	Objetivos	Situaciones decisivas	Necesidades y valores sociales	Cuestiones políticas	Cuestiones y problemas específicos
Finalidad	Relacionar los resultados con los objetivos	Proporcionar los conocimientos suficientes y una base valorativa para tomar y justificar decisiones	Juzgar los méritos relativos de bienes y servicios alternativos	Describir y valorar los costes y beneficios de las diversas opciones políticas. Destacar los aspectos positivos y negativos de un programa	Facilitar la comprensión de las actividades y su valoración en una teoría determinada desde distintas perspectivas
Procedencia de las cuestiones	Planificadores y responsables del programa	Los que toman decisiones, sus ayudantes y los evaluadores	La sociedad en general, los consumidores y el evaluador	Los legisladores, los grupos políticos y los grupos interesados	La comunidad y los grupos de clientes del programa y los expertos
Cuestiones principales	¿Qué estudiantes alcanzan sus objetivos y en qué grado?	¿Cómo debe ser planificado, ejecutado y reciclado un programa determinado para que promueva el crecimiento humano y el desarrollo a un coste razonable?	¿Cuál de los diversos objetos de consumo alternativos representa la mejor compra que pueda hacerse, dados sus costes, las necesidades de los usuarios y los valores de la sociedad en general?	¿Cuál de entre dos o más políticas distintas logrará con mayor éxito resultados válidos a un coste razonable? ¿Son más importantes los aspectos positivos que los negativos de un programa?	¿Cuál es la historia de un programa y cómo es juzgado por quienes están implicados en él y por los que son expertos en el área en el que se incluye el programa?
Metodología	Contraste entre resultados y los objetivos. Diseños experimentales y cuasiexperimentales	Inspecciones, valoración de las necesidades, grupos de recomendación, observaciones y métodos experimentales	Listas de control, valoración de necesidades, evaluación sin metas, planificación experimental y análisis de costes	La técnica Delphi, diseños experimentales y cuasiexperimentales, guiones, pronósticos y procedimientos judiciales	Estudio de casos, informes, sociogramas, observación participante, triangulación y diseños emergentes

Tabla II. Modelos de evaluación de programas. (García Ramos 1992: 469).

a lo que un programa concreto espera y necesita (Midkiff y Burke, 1987). Se impone la colaboración entre especialistas (de los contenidos evaluados, de metodólogos, de evaluadores...) y la combinación de modelos (proceso y producto, evaluaciones internas y externas...) y de diseños metodológicos (experimentales y naturalistas (Finney y Moos, 1989) para, en cada caso, ajustar coherentemente el producto (ya sea informes orales y escritos, memorias, planes de acción para la mejora...) y las conclusiones de la evaluación a su finalidad y uso práctico.

Como corolario final de este apartado dedicado a los modelos de evaluación, Colás (1993: 21) extrae las siguientes conclusiones acerca de las características de la evaluación de programas:

1. La evaluación de programas no debe entenderse de modo alguno como uniformidad con un modelo de evaluación, metodología a aplicar, procedimientos de análisis o toma de decisiones.
2. El concepto de evaluación marcará la elección de posibles modelos teóricos y enfoques metodológicos a seguir (aspectos que trataremos en los capítulos siguientes).
3. El proceso evaluativo, aunque obedece a un esquema identificable con un proceso científico, admite diversidad de enfoques metodológicos. Técnicas de recogida de datos y distintas tomas de decisiones.
4. El concepto de evaluación también será clave en la determinación de aspectos importantes a considerar en la evaluación, tales como, temáticas, tipos de decisión, modos de prácticas de evaluación, rol del evaluador, audiencia, funciones, etc.
5. Un proceso evaluativo no se puede concebir como una actividad básicamente técnica y neutral, de mera aplicación de procedimientos científicos, sino como un proceso sucesivo de toma de decisiones coherentes que se inicia la delimitación del modelo teórico del que se parte, elección del tipo de evaluación a seguir, papel del evaluador, finalidad, etc., hasta terminar en la recogida de datos más puntual y la elaboración del informe evaluativo.
6. El contexto y la posición del evaluador serán aspectos muy importantes en la selección del modelo teórico a seguir.
7. Toda evaluación de programas cubre unas determinadas *funciones* en el programa, las cuales han de ser delimitadas y definidas previamente al proceso evaluativo. No son fijas ni están predeterminadas. El evaluador debe tomar opciones y decisiones en este punto.

4. LAS ESTRATEGIAS EN LA EVALUACIÓN DE PROGRAMAS

Hemos analizado hasta ahora los modelos de evaluación desde la insuficiencia de sus *estrategias* para implicar a las personas evaluadas en procesos de mejora

y la repercusión que tiene en el bajo uso práctico de la evaluación. Ahora veremos algunas *estrategias* prometedoras para elevar el grado de uso de la evaluación para la mejora educativa.

En *primer lugar*, y de acuerdo con lo que señalamos con respecto a las características de los programas, habrá que tener en cuenta que *es la evaluación la que tiene que adaptarse a las características contextuales del programa y no al revés*. Lo más probable es que, cualquiera que sea el objeto de evaluación, será más fácil adaptar a él las ideas previas acerca del modelo de evaluación que hacerlo al revés. Ello significa que el diseño de la evaluación deberá ser flexible para adaptarse a los intereses cambiantes de la comunidad (Cronbach según Stufflebeam y Shinkfield, 1987). La flexibilidad, según el mismo autor, incluye la revisión periódica de la planificación y las *estrategias* a adoptar con el fin de prestar atención a las sorpresas y problemas consiguientes.

En *segundo lugar*, como estrategia para la implicación y complementando la flexibilidad de la planificación, la evaluación deberá incluir en la agenda de gestión de la evaluación, *tiempos de trabajo* a lo largo de todo el proceso, en los que el personal pueda trabajar sobre la evaluación y reflexionar en cada una de sus fases de desarrollo. Como hemos señalado, la finalidad de la implicación de los que desarrollan los programas es incrementar la utilización de los hallazgos de la evaluación con lo que aumenta de este modo la consiguiente validez de los hallazgos (Brandon, Lindberg y Wang, 1993: 420). Louks-Horsley y colaboradores (1987) añaden por su cuenta que la implicación debe darse desde el mismo diseño de la evaluación:

> *"La inversión de tiempo y energía con la audiencia en las etapas tempranas del diseño de evaluación, incrementa las posibilidades de que la evaluación establecerá una diferencia de que será atendida y de que los informes serán realmente una guía para la toma de decisiones"*. (Pág. 40).

Provus concluye que la máxima implicación del personal del programa *en cada paso del proceso* de evaluación implica el establecimiento de una relación continua entre el personal del programa y el personal de evaluación (promovida particularmente por el personal de evaluación) que dé como resultado una *comunicación continua*. Las *estrategias* de comunicación continua entre evaluación y objeto de evaluación se revela como la máxima preocupación que debe tener el evaluador y prácticamente la única tarea, entre todas las que se realizan en la evaluación, que debe ser de entera responsabilidad del evaluador. El papel del evaluador deberá ser juzgado principalmente de acuerdo con los medios y esfuerzos que ha puesto para establecer y mantener la comunicación constante con las personas e instituciones implicadas.

Crosby (1982, cit. por Loucks-Horsley *et al.*, 1987) sugiere seis áreas de implicación de los participantes como sistema de aprendizaje:

1. Participar en clarificar las metas del programa e indicadores del éxito del programa.
2. Participar en el diseño del estudio.
3. Participar en desarrollar métodos de recogida de información.
4. Participar en recoger y proporcionar información.
5. Participar analizando la información.
6. Participar en informar sobre el aprendizaje.

En *tercer lugar*, la medida que quizás sea la más importante, consiste en que las personas, grupos de interés y perspectivas involucradas en el objeto de evaluación sean las *que determinen los aspectos o dimensiones prioritarias a evaluar*. Esta idea es señalada por Duke y Corno (1981, citados por Louks-Horsley, 1987) quienes indican que los afectados por el programa deberían estar implicados en la formulación de las preguntas e interpretación de las respuestas de evaluación además de controlar el uso que se vaya a hacer de la información resultante. Para Provus, la participación de los implicados en la selección de prioridades de evaluación es tan importante que le parece la única garantía para que la evaluación sea aceptada. Señala que la única certeza para dicha aceptación es que los propósitos del programa sean definidos por el personal del programa y los métodos de cambio sean determinados por ellos también. Por su parte Cronbach alerta acerca del perjuicio que puede suponer para la evaluación el centrar prematuramente sus prioridades en problemas sociales complicados al margen de las necesidades prácticas.

La *cuarta cuestión* a tener en cuenta se refiere a la *relación de la evaluación con los distintos intereses y perspectivas distintas* acerca del objeto de evaluación. Stake ha puesto una enorme energía en subrayar que las distintas audiencias necesitan distinta información y que la evaluación deberá ofrecer esa información diferenciada para cada perspectiva (Stake, 1987). La perspectiva de Stake parece lógica, es decir, una vez reconocida la existencia de diferentes perspectivas es razonable la idea de reunir a todas ellas en un área común de encuentro/consenso al que por convención podemos acordar denominar verdad. A partir de la información diversa parece sugerirse la posibilidad de construir un puzzle común en el que sin perder las diversas perspectivas, todas ellas integren una descripción acerca de la realidad. Sin embargo, como hemos visto, a Stenhouse esto no le parece posible. Desde mi punto de vista el empeño de Stake acerca de describir las cosas como son (Stenhouse) puede ser desde la perspectiva de la evaluación para la mejora, un ejercicio que en la práctica resulte inoperante (aún en el caso de lograrse técnicamente), difícil por abstracto e irreal para los participantes.

Las diferencias entre perspectivas e intereses de las audiencias son de naturaleza diferente (gestores, patrocinadores, administradores, destinatarios de los programas...). Ello introduce problemas de poder y de condicionamiento tanto del programa como del proceso de su evaluación. Ante la diversidad de intereses es difícil mantener una práctica de la evaluación estrictamente instrumental como vehículo de información tal y como quiere Stake.

La evaluación, a partir de las finalidades que hemos señalado para ella, como instrumento para la mejora, debe identificar lo antes posible las intenciones de cada perspectiva y, en coherencia con los fines y usos elegidos, deberá decidir para quién es la evaluación, quién tomará las decisiones de su diseño y desarrollo. Ésta es la opinión de Loucks-Horsley *et al.* (1987: 165) cuando afirman que ésta es la decisión primera y más importante al diseñar una evaluación. Esta decisión afecta a todo lo demás que sigue. Los problemas que hemos visto acerca del uso y la comunicación entre evaluación y usuarios, señalan a los directamente responsables de desarrollar el programa o de hacer funcionar la institución como los actores principales de la evaluación para la mejora. Por eso es imprescindible que, como señalábamos en el punto anterior, sean los que decidan acerca de cuáles deben ser las cuestiones que debe abordar la evaluación. Un argumento más, acerca del punto de vista aquí defendido a favor de priorizar la implicación de los que desarrollan en la práctica los programas e instituciones evaluadas, que como ya señalamos al principio, no excluye otras posibles finalidades y usos de la evaluación, es el que nos sugieren Brandon, Lingberg y Wang (1993) cuando señalan que quienes toman decisiones en los programas y participan en las decisiones, tanto acerca de los programas como de las evaluaciones que sobre ellos se realizan; sin embargo, los que deberían participar, dicen los autores señalados, son los afectados por esos programas.

La anterior reflexión plantea el *quinto elemento a tener en cuenta* por la evaluación que nos proponemos. Se refiere a las *condiciones existentes en el contexto*, unas favorables y otras desfavorables, *para el desarrollo de una evaluación* para la mejora de la calidad educativa y para el desarrollo de cualquier programa. Sintetizamos cinco aspectos referidos al contexto y estrechamente relacionados, de los que la evaluación se tendrá que ocupar para modificarlos y, por otro lado, adaptar su diseño y desarrollo:

1) La voluntad del personal directamente implicado en la aplicación del programa o institución que se evalúa. El grado de voluntad para implicarse en un proceso de evaluación de estas características depende del *status* que el programa o la institución tenga y el grado alcanzado en el desarrollo de la autonomía. En relación con el personal afectado, habrá que tener en cuenta el nivel de desarrollo alcanzado como grupo en el conocimiento de sus funciones y la base teórica en la que fundamenta su acción profesional. De la discusión de estos aspectos puede elaborarse un plan de formación que documente las decisiones a tomar acerca del modelo, diseño y desarrollo de la evaluación (Simons, 1995).
2) Las condiciones externas que posibilitan o impiden la autonomía y la democracia. La capacidad y voluntad de decisión por parte de los implicados está determinada por la trama legislativa y la tendencia del sistema educativo (hacia la burocracia y jerarquización o hacia la democracia)

que condicionan la vida interna de centros, programas e instituciones en general. La necesidad de un marco plenamente democrático es señalada por House (1988). Al respecto, Brandon, Lindberg y Wang (1993) se preguntan cómo es posible implicar a los afectados por el programa evaluado en un marco no democrático y también: ¿Cómo puede ser desarrollada y mantenida la implicación democrática en programas en los que los beneficiarios no tienen autoridad para tomar decisiones?

3) Disposición y tiempo entre las múltiples tareas del personal. Los promotores de la evaluación, sean éstos internos o externos al objeto de evaluación deberán tener en cuenta que actualmente y en nuestro contexto educativo, la evaluación realizada de forma explícita y planificada es, si llega a realizarse, una parte muy pequeña entre las prioridades del personal y entre las múltiples tareas, casi todas más urgentes, que deben acometerse en un programa o en una institución. Sin embargo, la evaluación deberá lograr un tiempo y un espacio mínimos de dedicación entre el personal, sin el cual será imposible llevarla a cabo. En la medida en que la evaluación no ocupe un lugar en la preocupación del personal y la institución, será una tarea que por la presión de otras prioridades más inmediatas, acabará abandonándose. Por ello la evaluación si quiere tener una oportunidad deberá pensar en *estrategias* que le permitan abrir un espacio y un tiempo en el entramado de activismo del programa.

4) La complejidad del programa determina las estrategias de la evaluación. Brandon, Lindberg y Wang (1993) señalan que las *estrategias* para la implicación dependen del grado de complejidad y estructuración de los programas a evaluar. Un programa o institución complejo, en el que difícilmente vamos a encontrar una identidad entre lo que ocurre y lo que se planifica, que debe adaptarse a la situación cambiante sobre la que actúa y cuya actividad se ramifica en una gran diversidad de subprogramas, actividades y contextos muy diversos, modelará necesariamente, de forma muy singular, las *estrategias* para la implicación y su agenda de gestión serán igualmente complejas.

5) El último factor al que queremos hacer referencia como aspecto al que la evaluación para la mejora debe atender con la finalidad de elevar su grado de uso práctico, hace referencia a la necesidad de tener en cuenta las prácticas evaluativas preexistentes en el programa objeto de evaluación. Partimos de la idea de que el personal de los programas ya hace evaluación, en general de forma implícita y bajo diversas formas, como manera de regular su práctica y tomar decisiones tanto colectivas como individuales. Esas prácticas informales deben ser reconocidas en su contexto por parte de los evaluadores (internos o externos) antes de hacer propuestas acabadas de planes de evaluación. Creemos que una de las causas del fracaso

de las evaluaciones podría deberse a que, en cierto modo, las propuestas formales de evaluación chocan y entran en contradicción con las prácticas preexistentes.

Señalábamos anteriormente que para aumentar el grado de uso de la evaluación, los implicados en los programas debían participar en el diseño de la evaluación desde las primeras fases. De acuerdo con ese principio, uno de los aspectos importantes de la evaluación en el que deberá participar el personal del programa o institución a evaluar, es en determinar el grado de formalización de la evaluación que están dispuestos a asumir como colectivo, dadas las características contextuales en las que se desarrolla su trabajo (tiempo disponible, implicación, formación, autonomía para decidir, recursos, condicionantes externos...). Por ejemplo, es más oportuno que las propuestas acerca de los instrumentos a usar por el personal del programa (para lograr la implicación y para acceder a información difícil de obtener) se adapten a las limitaciones prácticas señaladas y a los sistemas de registro ya establecidos por ellos anteriormente (para a partir de ahí mejorarlos paulatinamente), que proponer sistemas de recogida de información sofisticados e inoperantes en ese contexto específico.

Por nuestra parte estamos de acuerdo con Stufflebeam y Shinkfield (1987) cuando, interpretando a Cronbach, afirman que la evaluación debe ser planificada sistemáticamente del mismo modo que se planifica una investigación. Sin embargo, Kemmis (1986: 123) nos advierte que la "planificación racional" de la evaluación puede tratar las evaluaciones de programas de una forma burocratizada que no hace justicia al carácter orgánico y reflexivo de la vida social y educativa y ése es el problema. Igualmente estamos de acuerdo con Stake (1987) cuando expresa la idea de que los evaluadores deben preocuparse por transformar la práctica de las evaluaciones informales en formales. Lo que aquí discutimos es el proceso que va de una práctica a otra y cuál es el papel de la propia evaluación en todo ese proceso.

Trasladar a la práctica *estrategias* como las que acabamos de explicar, significa para la evaluación que, además de ocuparse de las tareas que le son propias tradicionalmente (tales como establecer los objetivos, las dimensiones, los criterios, los indicadores, las técnicas de recolección de información, seleccionar y elaborar y/o adaptar instrumentos, analizar y elaborar conclusiones, etc.), deberá realizar actividades/operaciones, previamente planificadas en el diseño, en las que participen las personas directamente involucradas en los programas (profesionales que lo desarrollan en la práctica y destinatarios) con la finalidad, por un lado, de lograr su implicación final en el proceso de mejora resultante y, por otro lado, de establecer las condiciones contextuales, a las que nos hemos referido, que posibiliten la realización de la evaluación (dedicación de tiempo, medios humanos y materiales, etc.).

Las *estrategias* pueden tener un formato muy diferente (discusión para llegar a acuerdos, elaborar documentos, comunicar propuestas, formación, analizar, reu-

niones en pequeños y grandes grupos, etc.). Finalmente, el contenido de las *estrategias* dependerá de cada fase en la que se encuentre la evaluación. Por ejemplo, si se trata de establecer la finalidad de la evaluación quizá sea necesario hacerlo a través de la discusión de propuestas alternativas presentadas por escrito por parte de los evaluadores; si nos debemos de poner de acuerdo acerca de las características de la evaluación que llevaremos a cabo, será conveniente realizar alguna actividad puntual de formación acerca de los modelos de evaluación y sus ventajas; si se trata de determinar las dimensiones a evaluar podrá hacerse priorizando, a través de alguna técnica en gran grupo aquellos aspectos de más interés práctico, etc.

En resumen, la evaluación de programas incluye operaciones con tres finalidades: técnicas, modificar las condiciones del contexto para que la evaluación pueda realizarse y de implicación de las audiencias en todo el proceso. Todo ello deberá tenerse en cuenta en el diseño de evaluación que abordamos a continuación.

5. ELEMENTOS Y ESTRUCTURA DEL DISEÑO DE PROYECTOS DE EVALUACIÓN DE PROGRAMAS

A pesar de que, como ocurre también con los programas, es común que haya diferencias entre la evaluación diseñada y la evaluación finalmente realizada, debido a los cambios necesarios que hay que introducir en el diseño para que se adapte al proceso, es necesario elaborar la propuesta de evaluación lo más ajustada y realista posible como si de un diseño de investigación se tratara. El documento resultante es un instrumento de trabajo imprescindible para que, desde el principio, las distintas personas implicadas lo tomen como referencia de actuación. La primera propuesta de proyecto de evaluación, indudablemente es de los evaluadores. Lo que estamos proponiendo es que en el proyecto se incluyan *estrategias* para que el proyecto llegue a contar con el mayor número de apoyos posibles dentro del programa, por lo que, una vez revisado con las personas implicadas, seguramente sufrirá cambios necesarios, los cuales harán que pase a ser un proyecto compartido. A su vez, para diseñar el proyecto inicial, el evaluador deberá conocer el programa familiarizándose con él (Alvira, 1991: 22).

Distinguimos tres fases o partes que vamos a explicar a continuación. Las fases son:

1. Cuestiones previas a la elaboración del diseño/proyecto de evaluación.
2. Explicar las características de la evaluación por la que se opta.
3. Diseño de la evaluación.

La tabla III que mostramos a en la página siguiente sintetiza las cuestiones previas relativas a la elaboración de un diseño de evaluación.

En esta primera fase debemos hacer una propuesta de finalidades de la evaluación, del papel que jugarán los evaluadores y de un análisis del objeto/programa a evaluar. En esta fase incluimos también, como hemos explicado en el apartado anterior, las *estrategias* de evaluación que prevemos realizar en esta fase.

ELEMENTOS Y ESTRUCTURA
PARA LA ELABORACIÓN DE PROYECTOS DE EVALUACIÓN DE PROGRAMAS (I)

FASE A: **TRES CUESTIONES PREVIAS A LA ELABORACIÓN DE UN DISEÑO DE EVALUACIÓN**

Aspectos a aclarar, concretar, explicitar en un **proceso** previo al diseño y posterior desarrollo de la evaluación.

1. Quién o quienes demandan la evaluación. Para qué se quiere realizar la evaluación; cuál es el objetivo que se pretende con la evaluación; para qué evaluar; CUÁL SERÁ EL USO, LA FINALIDAD DE LA EVALUACIÓN.

– Informe técnico	– Aportar argumentos para el debate crítico
– Conocer	– Tomar decisiones
– Adecuación al mandato legal	– Rendir cuentas.
– Mejorar	– Etcétera.

2. Relación del o los evaluadores con la institución o programa a evaluar; PAPEL DE LOS EVALUADORES CON RESPECTO AL PROGRAMA
   - Evaluador externo (consultor, experto técnico, apoyo a procesos de evaluación y mejora...).
   - El evaluador/evaluadores es un equipo interno al propio objeto de evaluación.
   - La conexión con el objeto de evaluación se hace a través de un reducido grupo miembros de la institución o vinculados con el programa a evaluar. Caracterización de ese grupo: responsables, directivos, elegidos...
   - La relación se hace directamente con todos los miembros de la institución o programa a evaluar.
   - ETC.
3. ANÁLISIS DEL PROGRAMA A EVALUAR. Naturaleza del objeto de evaluación. Qué se quiere evaluar, qué aspecto o enfoque de lo que se quiere evaluar es lo que interesa por su relevancia para el programa, etc.
   EJEMPLOS:
   - Características relevantes.
   - Finalidades más importantes del programa.
   - Procesos que se ponen en marcha.
   - Qué estructura tiene.
   - Etcétera..
4. ESTRATEGIAS DE EVALUACIÓN PARA LA FASE **A.**

Tabla III. Esquema de elementos de estructura para diseñar proyectos de evaluación de programas (parte I).

En cuanto a la ***finalidad de la evaluación***, se trata de establecer *para qué* se pretende hacer la evaluación y *cuál será su uso* práctico. Ambos aspectos determinarán el resto del diseño. El evaluador deberá indagar acerca de las razones por las que se demanda la evaluación. Para ello atenderá a todas las personas implicadas en el programa (audiencias). Frecuentemente nos encontraremos con que la finalidad de la evaluación, lo que se espera de ella y su uso es distinto para cada una de las audiencias.

Las finalidades de la evaluación pueden ser muy diversas tal como se recoge en el esquema precedente. Para que la realización de la evaluación se ajuste adecua-

damente a lo que se espera de ella, es necesario distinguir desde el principio cuáles son las finalidades y usos formalmente explicitados y cuáles son las finalidades realmente sentidas por las distintas audiencias (objetivos ocultos). Puede haber diferencia entre unas y otras. Hay que tener presente que finalmente se juzgará a la evaluación por el grado de satisfacción de las finalidades prácticas más que por las formales.

Lo ideal es que la finalidad o finalidades de la evaluación se expliciten en forma de objetivos aunque al hacerlo de ese modo corremos el riesgo de no lograr expresar de forma realista, clara y práctica el cometido de la evaluación. Además, un mal uso de los objetivos previamente establecidos puede hipotecar el desarrollo de la evaluación si lo condicionamos al logro de los mismos. Independientemente de la forma en que se establezcan las finalidades, hemos señalado antes que el diseño de evaluación es una hipótesis de trabajo que debe permanecer abierto a los cambios que juzguemos necesarios introducir durante el proceso.

La idea de «uso de la evaluación» trata de enfatizar la necesidad de establecer cómo, quién, para qué se hará la evaluación: a quién beneficiará, quién controlará la información resultante y sus conclusiones, cómo participaran en ello las distintas audiencias, etc.

El segundo de los puntos de esta fase se refiere al ***papel del evaluador*** o evaluadores. Las funciones del evaluador o evaluadores no quedan condicionadas porque éstos sean internos o externos al objeto evaluado. Lo importante es que el papel de los evaluadores se ajuste a la finalidad de la evaluación. En este sentido el carácter interno o externo de la evaluación está más determinado porque la prioridad (finalidad) de la evaluación sea interna o externa que porque los evaluadores pertenezcan profesionalmente al programa evaluado o sean externos a él. Dependiendo de que el promotor de la evaluación (el que la demanda, la subvenciona, la controla, la impulsa, etc.) sea interno al programa o externo a él, la evaluación puede tener prioridades internas o externas porque la finalidad de la evaluación sí está determinada por los intereses de quienes la promueven. Así, seguramente, a autoridad administrativa de un departamento que promueve un programa social le puede ser más urgente tener datos acerca del impacto del programa mientras que al personal que lo desarrolla en la práctica le puede interesar más que la evaluación sea instrumento de apoyo para la mejora desde la comprensión de la práctica, etc.

Es necesario que, por parte del evaluador y desde el principio, se haga una propuesta acerca de su papel y de las tareas que asumirá y no asumirá. Algunos de los posibles perfiles se incluyen en el esquema precedente. En cualquier caso, sobre el evaluador recae la ejecución de la parte técnica de la evaluación. Sin embargo lo que queda por establecer es la amplitud de la implicación y colaboración de las distintas audiencias en todo el proceso de evaluación para que sus conclusiones finalmente se lleven a cabo. Para lograr tal fin, lo ideal es que se establezcan equipos mixtos entre el personal interno (comisión de evaluación interna) y el externo

(evaluación mixta interna-externa; Tiana, 1998), entre expertos en evaluación y profesionales con experiencia en el desarrollo de programas como el que se evalúa. Todo ello es lo que se debe explicitar en este punto como propuesta para posteriormente ser revisada por las distintas audiencias.

El tercer aspecto de esta primera fase se refiere al ***análisis del programa*** a evaluar. Las finalidades del diseño de este paso son al menos dos: a) llegar a establecer puntos de vista comunes (compartidos) acerca del programa que se quiere evaluar, entre los evaluadores y las personas del programa; y b) elaborar un marco de referencia (criterios) teórico y práctico, lo más compartido posible, desde el que realizar el análisis del programa evaluado.

Para ello el evaluador deberá anticipar una propuesta, que deberá ser revisada junto a las personas directamente implicadas, en la que a) muestre su conocimiento del programa que pretende evaluar: objetivos, prioridades, fases, metodologías, dificultades prácticas, etc.; b) identifique, según su propio criterio, qué aspectos del programa son los fundamentales para sus patrocinadores, los destinatarios del mismo, etc; y c) relacione los aspectos más importantes del programa con el conocimiento teórico y experiencias existentes, a partir de los que elaborar marcos interpretativos (criterios). Si, por ejemplo, el programa trata de introducir mejoras en la organización institucional, el evaluador deberá proponer la perspectiva organizativa, la teoría y algunas experiencias desde las que analizar las mejoras organizativas; si se trata de una metodología determinada, el evaluador deberá aportar elaboraciones teóricas desde las que extraer los criterios a partir de los que analizar la metodología; o si, en fin, se trata de un programa relacionado con la formación del profesorado, el evaluador aportará los modelos de formación y las características por las que se distingue en la práctica cada uno de ellos. Un programa, por sencillo que sea, incluirá todos estos temas y/o algunos otros. Al respecto es ilustrativo el trabajo realizado por De Miguel (1995), para tal fin, en el que identifica y desarrolla la teoría de cada una de las variables institucionales relacionadas con la calidad: estructura de la organización, calidad del currículum, calidad de la enseñanza, la innovación, etc.

Posiblemente uno de los mejores productos finales de una evaluación realista, atenta a las dificultades prácticas de la mejora de la calidad, sea la elaboración de un marco conceptual compartido por las audiencias a partir del que pueda interpretarse, de forma unificada, el proceso y el producto del programa evaluado.

Una vez que el evaluador analice el programa, haga observaciones directas del programa, realice entrevistas con el personal y los destinatarios para familiarizarse con él, estudie otras experiencias, etc. estará en disposición de hacer una propuesta que incluya los tres puntos anteriormente mencionados e igualmente estará en disposición de comenzar a revisar la propuesta de evaluación con las audiencias. La planificación de la forma en que se hará la revisión de la propuesta es lo que hemos llamado *estrategias* de evaluación.

Cada uno de los puntos de esta *fase A* **deben incluir estrategias** en las que se explique qué queremos hacer, cómo lo queremos hacer y con quién para revisar las propuestas realizadas acerca de los tres puntos incluidos en esta primera fase: la finalidad de la evaluación, el papel de los evaluadores y el análisis del programa a evaluar.

ELEMENTOS Y ESTRUCTURA
PARA LA ELABORACIÓN DE PROYECTOS DE EVALUACIÓN DE PROGRAMAS (II)

FASE B: **EXPLICAR LAS CARACTERÍSTICAS DE LA EVALUACIÓN POR LA QUE SE OPTA.**

- Evaluación por objetivos.
- Evaluación experimental, cuasiexperimental.
- Evaluación holística, iluminativa, respondiente, (deliberativa, interpretativa... -interés práctico-).
- Autoevaluación, evaluación democrática.
- Evaluación sociocrítica -interés emancipador- etcétera.

Opción que depende de:

1. Las finalidades asignadas a la evaluación.
2. La consideración que tengamos sobre el objeto a evaluar y las características de la organización en la que se desarrolla.
3. Circunstancias que caracterizan y condicionan el programa (económicas, normativas/legales, etc.).

La opción que se tome acerca de las características de la evaluación a realizar, determinará todos los elementos, fases, *estrategias* y técnicas de la evaluación.

–ESTRATEGIAS DE EVALUACIÓN PARA LA FASE **B**.

Tabla IV. Esquema de elementos y estructura para diseñar proyectos de evaluación de programas (parte II).

Una vez que en la fase A hemos resuelto los aspectos prácticos acerca de la evaluación, como son las finalidades que se le asignan por parte de las personas que la demandan, el papel del evaluador y el análisis de los problemas y prioridades del programa a evaluar, abordamos en esta fase la *fundamentación teórica de lo que pretendemos hacer*. No se trata de partir de esquemas teóricos para adaptar la realidad a ellos sino que al contrario, las demandas y condiciones de la práctica deberá ayudarnos a elaborar las características teóricas de la evaluación atendiendo a las finalidades ya propuestas en la fase anterior, más que al dogmatismo de los modelos. Los modelos de evaluación no se ven ya como excluyentes y contradictorios entre sí, no se trata de intercambiar unos modelos de evaluación por otros sino de complementarlos, elaborando síntesis creativas que se adecúen a las necesidades que se nos demandan desde el programa. En este sentido, el programa lo es todo y la evaluación es un instrumento eficaz para su mejora.

Por lo tanto, es importante que desde el principio se haga una propuesta global acerca de cuáles son las características generales de la evaluación, cuáles

son sus fases teóricas y qué diferencias existen con respecto a otras posibles propuestas de evaluación. A partir de esa propuesta, y a través de las *estrategias* de evaluación, se discutirá con las personas implicadas lo que se puede esperar de la evaluación que se está proponiendo y qué no se debe esperar de la propuesta.

Como venimos señalando, es conveniente que diseñemos *estrategias* para discutir todo ello y evitar confusiones acerca de lo que será y no será la evaluación. Una estrategia conveniente para esta fase parece ser el proponer una planificación de formación puntual para todas las personas afectadas, sobre modelos y experiencias realizadas de evaluación de programas similares a los que se pretende evaluar (Simons, 1995). De tal manera que las personas implicadas tengan argumentos para modificar o, en su caso, rechazar la propuesta que les hacemos acerca de las características de la evaluación.

En la *fase C* incluimos los elementos propios del diseño de evaluación. Comenzamos proponiendo los medios y recursos con que deberá contar la evaluación que pretendemos hacer para que ésta pueda desarrollarse. Lo importante es lograr que la evaluación no quede al margen del programa y que el personal asuma la presencia constante de la evaluación. De ello depende que la evaluación ocupe un espacio propio entre las operaciones del programa de manera que, a la vez que se realiza, se considere como un aspecto imprescindible por parte de las audiencias. Lograrlo no es siempre fácil, es desde la evaluación desde donde se tiene que hacer el esfuerzo para conseguir el espacio necesario entre el entramado de tareas y actividades regulares del programa. De lo contrario, la evaluación se puede convertir en un residuo marginal, que no es atendido por nadie y que es sólo una preocupación particular de los evaluadores.

Para ello, desde el principio deberá quedar claro cuáles son las necesidades de la evaluación y, sobre todo, qué estructura organizativa se va a adoptar para atender a su realización: comisiones internas, coordinaciones, etc.

El segundo elemento de esta fase se refiere a las ***dimensiones*** de la evaluación. Una vez que en un paso anterior de la fase A se ha analizado el programa de forma independiente, por parte de los evaluadores, éstos están en disposición de proponer sobre qué elementos del programa deberá focalizarse la evaluación para hacer su seguimiento, análisis y valoración. Por lo tanto las dimensiones a evaluar tienen que captar lo que de esencial aporta el programa: el significado del programa, lo que pretende; los objetivos; la estructura, fases y elementos claves del programa; los individuos implicados; las acciones; identificar los impedimentos que tiene para desarrollarse, los recursos de todo tipo de que dispone como apoyo para su desarrollo y las consecuencias esperadas del mismo.

Igual que ocurre con el análisis del programa que, como propuesta hacen los evaluadores, como vimos anteriormente, en el caso de las dimensiones los evaluadores hacen también una propuesta independiente que posteriormente, a través de las *estrategias* de evaluación deberán contrastar con el punto de vista de las audien-

ELEMENTOS Y ESTRUCTURA
PARA LA ELABORACIÓN DE PROYECTOS DE EVALUACIÓN DE PROGRAMAS (III)
FASE C: **DISEÑO DE LA EVALUACIÓN**

1. Establecer los **medios y recursos** necesarios para realizar la evaluación que estamos proponiendo:
   a) Financiación.
   b) Espacios.
   c) Inversión en tiempo por parte de la institución o programa evaluado para participar en el proceso de evaluación.
   d) Materiales a utilizar: fotocopias, bibliografía.
   e) Cursos de formación inicial, durante el proceso...
   f) Negociar líneas de actuación encaminadas a **establecer un tejido de colaboración/implicación** en todo el proceso de evaluación: grupo de contacto, grupos de trabajo, equipos de autoevaluación, elección de prioridades, etc... Tener en cuenta que este último aspecto deberá figurar en la agenda de trabajo.
   g) ETC.
2. Establecer los distintos aspectos o **DIMENSIONES** del objeto de evaluación. Para ello habrá que partir del análisis realizado en el punto 3 del apartado A de este diseño.
3. Explicitar indicadores para cada uno de los distintos aspectos o dimensiones.
4. Decidir acerca de la estructura de la recolección de datos y evidencias:
   - si la evaluación se centrará o no en el análisis de
     - INTENCIONES.
     - REALIZACIÓN OBSERVADA.
   - en qué momentos del programa se realizará la evaluación:
     - de CONTEXTO,
     - de ENTRADA,
     - de PROCESO,
     - de PRODUCTO.
   - si existirá una evaluación INTERNA y una evaluación EXTERNA o solo una de las dos.
5. Establecer las distintas técnicas para la recogida de los datos, evidencias e información necesarios para realizar la evaluación. Para cada uno de los indicadores establecidos elegir la técnica o técnicas que se utilizaran.
6. Elaborar una agenda para aplicar el diseño en sus dos vertientes: metodológica y de negociación o proceso.

Una vez que se lleve a cabo el plan de evaluación qué previsión hacemos con respecto a:

7. ¿Cómo se ORGANIZARÁN Y PRESENTARÁN los datos y evidencias recogidos acerca de cada una de las dimensiones de evaluación?, ¿Quién se encargará de organizar los datos?, ¿A quién se presentarán los datos?, ¿De qué modo: oralmente, a través de documentos y medios audiovisuales...?
8. Análisis y discusión para la
9. Elaboración de planes de acción, elaboración de informes de progreso o memorias, la elaboración de un informe técnico, etc..., según sea el modelo de evaluación por el que se ha optado.
10. METAEVALUACIÓN.
11. ESTRATEGIAS DE EVALUACIÓN PARA CADA UNO DE LOS PUNTOS DE LA FASE C

Tabla V. Esquema de elementos y estructura para diseñar proyectos de evaluación de programas. Parte III.

cias. En este sentido es importante que los evaluadores aporten su propia visión del programa y amplíen las perspectivas acerca del mismo, siempre como propuesta inicial. La ampliación de perspectivas, en aquellos programas que lo necesiten puede ser una aportación enriquecedora desde la evaluación. Por ejemplo, puede que un determinado programa no esté teniendo en cuenta la participación y atención de algún sector importante o puede que no se tenga en cuenta algún factor como la formación de los técnicos, la organización institucional, los procesos de toma de decisiones, etc. Si es así, y se considera importante, habrá que incluirlo entre las dimensiones a analizar.

Por otro lado es recomendable que el sistema de dimensiones y subdimensiones resultante (que siempre suele ser un sistema complejo y difícil de establecer, dependiendo del programa) recorra el programa desde sus aspectos más amplios y estructurales (*macro*), tales como la política educativa o las estructuras de decisión externas al programa que condicionan positiva o negativamente su desarrollo, hasta sus aspectos más concretos (*micro*) tales como la metodología de aprendizaje o tratamiento, la estructura de las actividades individualizadas, etc. Un ejemplo de sistema de dimensiones y subdimensiones con las características como las que nos referimos, es el elaborado en la evaluación de los Centros del Profesorado (Gobantes, 2000).

Colás Bravo (1993: 58) sugiere la utilización del cubo de Hammond (1973), representado en la figura 1, en el que se relacionan cuatro referentes: las fases y los elementos clave del programa, los individuos implicados en el programa a evaluar, los objetivos del programa y las acciones en las que se involucran el programa, los objetivos y los agentes.

El tercer elemento de esta fase son los ***indicadores*** de la evaluación. Si hemos realizado el análisis del programa, tal como indicábamos en el apartado A3, habremos identificado las características teóricas de aquellos elementos importantes del mismo. De ese modo, por ejemplo, sabremos cuáles son las referencias teóricas y las características distintivas de la metodología que se pretende, de un proceso de resolución de conflictos, de un sistema de organización para la toma de decisiones dentro del programa, de las finalidades del mismo, etc. A partir de ahí, ahora se trata de establecer *para cada una de las dimensiones y subdimensiones* ya explicitadas en el paso anterior, aquellos trozos de la realidad observada (cualitativa o cuantitativa) que nos alertan o nos (de)muestran la existencia de una característica esperada y producida por el programa (en el proceso o en el producto).

Las siguientes características que para Lázaro Martínez (1992) tienen los indicadores de evaluación nos permitirá perfilar el concepto de indicador de acuerdo con los fines de nuestra idea de evaluación para la mejora:

- Un indicador señala un aspecto esencial o accidental de una cosa, de forma que nos permite conocerla mejor y más profundamente.
- El indicador es una señal, una guía, ya que viene a establecer las referencias o el camino a seguir para llegar al conocimiento de algo.

DIMENSIONES Y SU CONTENIDO

1. MODELOS DE FORMACIÓN.
   * Asesoramiento a centros (individual y por parejas).
   * Atención a agrupaciones.
   * Actividades de formación (antigua nº 3): tipología, diversidad.
   * Autoformación.
   * Coordinación con otros servicios: inspección, orientadores, dirección general, STOEP.
2. ORGANIZACIÓN INTERNA DEL CEP.
   * Infraestructura de trabajo:
     uso del teléfono
     uso del ordenador
     personal
     kilometraje.
   * Organización de reuniones.
   * Coordinación del equipo pedagógico: etapas y niveles.
   * Funcionamiento de los órganos del CEP - democratización-: Consejo General, Consejo de Dirección, Equipo pedagógico.
   * Clima: relaciones humanas.
   * Circulación de la información dentro del CEP:
     Origen.
     Acceso democrático a la información.
   * Comisiones.
   * Presupuesto de funcionamiento.
   * Gestión económica.
3. OFERTA/DEMANDA.
4. RECURSOS Y SERVICIOS OFERTADOS AL PROFESORADO POR EL CEP.
   * Servicios a los que nos referimos:
     Bibliografía.
     Laboratorio.
     Audiovisuales.
     Reprografía.
     Lugares de esparcimiento: café... charla informal...
     Otros...
   * Funcionamiento de los servicios.
   * Uso.
   * Características.
   * Objetivos.
   * Planificación.
   * Personal.
   * Actualización de fondos y recursos.
5. COORDINACIÓN CON OTROS SERVICIOS -GRADO DE AUTONOMÍA DEL CEP.
   * Niveles de participación del profesorado.

Tabla VI. Ejemplo de dimensiones y subdimensiones con su contenido

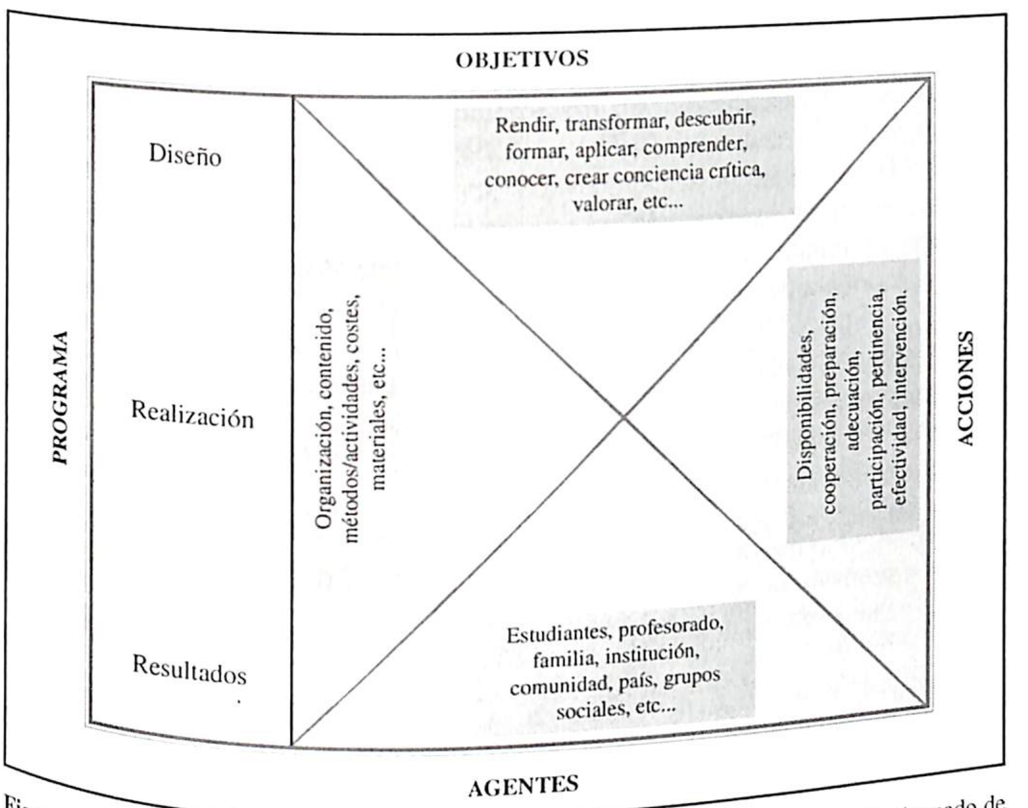


Figura 1. Cubo de Hammond (1973), su finalidad es generar cuestiones de evaluación (tomado de Colás Bravo (1993: 59).

- El indicador ha de referirse a una situación concreta, esto es, el fenómeno estudiado por el indicador tiene un espacio y un tiempo determinados.
- Todo indicador posee un carácter parcial de las manifestaciones del objeto a estudiar. Los datos que suministre deben ser analizados conjuntamente con el resto de indicadores.
- Todo indicador participa de un sentido secuencial. En el análisis de cualquier fenómeno se busca una cierta presencia en el presente que puede tener sus raíces en una situación pasada. En nuestro caso, el análisis del pasado y el devenir del proceso es objeto fundamental de estudio.
- El indicador tiene un sentido de anticipación diagnóstica, ya que intenta localizar síntomas de la manifestación de algo, de forma que un conjunto de los datos puede servir de conocimiento previo de algo, que se va enriqueciendo a medida que se acumula más información.
- Un indicador debe recoger la información precisa para facilitar la comparación entre lo manifestado y lo establecido como deseable, aunque con la cláusula de que deben ser congruentes con el contexto general de análisis del fenómeno.

Finalmente, Landsheere (1994) cita las indicaciones que Darling-Hammonn (1991) propone para la utilización constructiva de los indicadores:

1. Las cifras no bastan. Un conocimiento de la enseñanza, del aprendizaje y del funcionamiento de las instituciones educativas debe permitir a todas las partes interesadas interpretarlas correctamente. No porque las cifras sean precisas son válidas.
2. Solamente una multiplicidad de indicadores puede dar una imagen correcta de una situación. Los aspectos de la educación que constituyen el blanco directo de la política educativa no son los únicos interesantes. Importa, particularmente, tener en cuenta las características duraderas del sistema educativo.
3. Los sistemas de indicadores deben ser redundantes y han de revisarse continuamente. El peligro de corrupción de un indicador se reduce si se procede mediante múltiples medidas de un mismo aspecto del sistema.
4. Haya o no indicadores, una política debe estar basada en una gran comprensión de la naturaleza de las motivaciones individuales y de los cambios organizativos, así como de la influencia de dichos factores sobre los rendimientos.
5. Los indicadores deben ser utilizados como el agua en el molino de la evaluación de una situación, y no como hechos suficientes para poner en marcha acciones que podrían ser ineficaces. El interés de los indicadores estriba en que proporcionan señales a propósito de cuestiones importantes y de problemas que podrían plantearse. En estos casos se impone un profundo examen.

Los indicadores deben ser lo más concretos posible. Hacer este trabajo desde el diseño facilitará la tarea posterior de organización de la recolección de la información a partir de ellos.

La información que se recogerá, en cualquier tipo de programa, es, según la fase, de contexto, de entrada (*input*) de proceso y de producto (Stufflebeam, 1987), sobre las intenciones y la realización observada (Stake, 1983).

Una vez elegida de la forma más precisa posible, la información que se necesita recopilar, el siguiente paso consiste en establecer las técnicas o métodos a través de los que se recogerá dicha información. Posteriormente, una vez discutido y revisado el diseño, se pasará a elaborar los instrumentos de nueva creación y a seleccionar y/o adaptar los ya existentes. Evidentemente la utilización de diferentes técnicas que se complementen, no debe ser una elección previa a la especificación de la información necesaria. Ello quiere decir que se deberán utilizar los métodos que mejor capten la información que necesitamos y que sean complementarios (triangulación metodológica), ya sean cuantitativos, cualitativos o ambos. No es objetivo de este capítulo desarrollar las metodologías de investigación ni las técnicas e instrumentos de recogida de datos. Son numerosos los manuales existentes al respecto, a los que remitimos al lector (por ejemplo: Goetz y LeCompte, 1988; Alvira Martín,

1991; Arnal, Del Rincón y Latorre, 1992; Colás Bravo y Buendía Eisman, 1992; Balcells, 1994; Del Rincón, Arnal, Latorre y Sanz, 1995; Cea D'Ancona, 1996; Rodríguez, Gil y García, 1996; Balles, 1997; Buendía, Colás y Hernández Pina, 1997; etc.). Con la secuencia del diseño de dimensiones, indicadores y técnicas/métodos de regida de datos de la evaluación es conveniente elaborar una tabla que represente su relación y que sirva de instrumento de discusión con las audiencias acerca de las pretensiones de la evaluación, tal como se representa en la fig. 2.

Tópico objeto de evaluación	Sistema de dimensiones y subdimensiones	Indicadores	Técnicas/métodos de recogida de datos y evidencias

Figura 2. Sistema de dimensiones-indicadores-técnicas.

El documento resultante sintetiza de forma muy gráfica la base metodológica de la evaluación que se propone; a los evaluadores les facilita la comunicación de la propuesta y a las personas implicadas en el programa les permite introducir los cambios y mejoras que consideren necesarias en el diseño.

En la evaluación del Plan Municipal de Educación y Atención a la Sexualidad Juvenil de Los Realejos (Cabrera, Gobantes, *et al.*, 2000), la dimensión acerca de la formación recibida por los profesionales difusores del programa, se organizó en diferentes subdimensiones referidas a: la formación en torno a contenidos sobre sexualidad, sobre el plan, sobre las actitudes acerca de la sexualidad, la satisfacción acerca de la formación recibida, la organización, etc. A título de ejemplo, recogemos en la tabla nº 7 los indicadores que se establecieron en el diseño acerca del conocimiento del plan.

Como se puede observar, entre los indicadores hay una cierta variedad en su nivel de concreción. Podemos considerar un contínuum en cuyos extremos se sitúa, por un lado, el nivel más general de concreción de las dimensiones y, por otro, la concreción máxima de los ítemes de, por ejemplo, un cuestionario o una entrevista. Entre ambos se mueven las subdimensiones y los indicadores más o menos concretos, más o menos próximos a uno y otro extremo. Como señalábamos anteriormente, cuanto más se aproxime la concreción de los indicadores a los de ítems, mejor organizado y dirigido estará el trabajo posterior de elección de técnicas, elaboración de instrumentos y recogida de información.

Como hemos señalado, con ello podríamos elaborar una representación gráfica en forma de tabla (figura nº 3) que exponemos en la página 121. En ella se relacionarían cada una de las dimensiones de evaluación con sus correspondientes cuestiones de evaluación. Para cada una de las cuestiones de evaluación se decidiría, a partir de los recursos existentes en la situación concreta de evaluación, la técnica o técnicas más adecuadas y eficaces dependiendo del tipo de información que demanda cada una de las cuestiones.

DIMENSIÓN	Subdimensión	INDICADORES
Formación para la realización del Plan	Conocimiento del Plan	Conocimiento de la implantación del Plan de Atención a la Sexualidad Juvenil que se desarrolla en Los Realejos Conocimiento de los sectores profesionales implicados en el Plan Participación en la elaboración del Plan Se conoce el Plan en su ámbito laboral, familiar o de amigos Necesidad de implantación del Plan Conocimiento de las acciones que se van a llevar a cabo dentro del Plan Conocimiento de los recursos que se van a poner en marcha para atender a la educación afectivo-sexual de los adolescentes y jóvenes Ha oído hablar del plan de Atención a la Sexualidad Juvenil de Los Realejos A quién va dirigido el Plan Se habla en tu ámbito del Plan Participación en las acciones del Plan Desarrollo adecuado del Plan Difusión adecuada del Plan Implicación por igual de todos los sectores Crees que la población de Los Realejos está suficientemente sensibilizada para participar en este Plan Conocimiento de acciones que no se están realizando de la manera adecuada Conocimiento de las acciones realizadas desde el Plan Colaboración en las acciones del Plan Confianza en las posibilidades de implantación práctica del Plan.

Tabla VII. Ejemplo de una subdimensión y sus indicadores en la planificación de la evaluación (Cabrera, Gobantes, et al., 2000).

En la tabla así elaborada, cada cuestión de evaluación estaría relacionada expresamente a uno o más instrumentos de tal manera que quedaría justificada su uso por la información que deberá captar. A partir de ahí podría sofisticarse la tabla añadiendo, por ejemplo, las fuentes de información o componentes de una agenda de gestión de la evaluación como tareas a realizar, responsables de cada tarea, fechas aproximadas, etc., tal como aparece en la figura 3. La finalidad de esta elaboración es reducir a una representación simplificada la complejidad de elementos a tener en cuenta en la evaluación y, por otro lado permite la comunicación y la discusión entre las personas implicadas en el proceso.

Sin embargo, para no complicar excesivamente la representación resultante, somos partidarios de elaborar una ***agenda*** esquemática de la evaluación, aparte de

Pregunta de evaluación	*Fuentes de información*	*Métodos de recolección*	*Responsabilidad*	*Tiempo*
¿Cómo afecta el entrenamiento y seguimiento de la conducta en el aula de los participantes?	*Profesores, supervisores, estudiantes.*	*Observaciones utilizando su comprobación* *Entrevistas*	*Profesor como cuadro de investigadores*	*Oct-Nov. (antes del entrenamiento) Abril-Mayo (después del entrenamiento)*
¿Han sido relevantes para las necesidades de los profesores las ofertas de desarrollo personal de este año? *¿Qué cambios se desean para el próximo año?*	*Profesores*	*Cuestionario escrito*	*Coordinador de desarrollo personal*	*Mayo*
¿Cuál es el impacto de la creación de equipos de desarrollo de personas basadas en la escuela?	*Profesores, directores, coordinadores y administradores de distrito. Miembros de la Comunidad*	*Etnografía*	*Estudiantes, graduados de la Universidad supervisados por el coordinador de evaluación de distrito.*	*Sept.-Agosto*

Figura 3. Ejemplo de entradas en una hoja de trabajo para la recogida de información (Hruska y Hawley, 1982).

la tabla de dimensiones-indicadores-técnicas. La agenda contendrá una propuesta, para su posterior discusión con las personas implicadas, que relacione, al menos: *a)* las tareas a realizar (actividades propias de la evaluación y *estrategias* para desarrollarla); *b)* los profesionales responsables de llevarlas a cabo (tanto evaluadores como del programa evaluado); y *c)* una previsión de su temporalización.

El diseño de la evaluación para los pasos a llevar a cabo una vez se realice la investigación (a partir del punto 7 de la fase C del esquema) debe prever tres problemas distintos:

a) El análisis, por parte de los evaluadores, de los datos y evidencias recopiladas.
b) El análisis y conclusiones por parte de las personas implicadas en el proceso de evaluación.
c) La elaboración del producto final de la evaluación (planes de mejora, informes, memorias, etc.).

En esta fase, más que en ninguna otra, se evidencia el necesario reparto de papeles entre los evaluadores y las personas implicadas en el programa que se evalúa. Mientras que los evaluadores, como expertos, deberán analizar la información recopilada, aplicando el tratamiento cualitativo o cuantitativo que les corresponda, según las técnicas seleccionadas, las audiencias deberán retomar el análisis de los

expertos para, en uno u otro grado de elaboración, tomar decisiones a la luz de los mismos.

La participación de las audiencias puede adoptar grados muy diversos: desde la recepción formal de los resultados de la evaluación, por parte de los responsables del programa (los que tienen poder para tomar decisiones acerca del futuro del mismo), totalmente elaborados en un informe los cuales se reservan el uso posterior de los mismos, hasta la participación en el análisis y la elaboración de conclusiones a partir de la información sistematizada por los evaluadores acerca de lo que ha ocurrido con el programa.

La participación, a la que nos referimos, de las distintas audiencias se realiza a través de lo que venimos denominando *estrategias de evaluación*, que deberán ser planificadas por los evaluadores. De las *estrategias* que se diseñen y lleven a cabo depende el grado de implicación de las audiencias en las conclusiones a las que llegue la evaluación y en la aplicación a la práctica de las mejoras del programa que finalmente se elaboren. Las *estrategias* para la implicación en esta fase pueden ser muy diversas: comunicación oral de los resultados de la evaluación, uso de medios audiovisuales, equipos de trabajo para la discusión y elaboración de alternativas, jornadas de evaluación en las que se planifique la participación de amplios sectores sociales y el análisis y elaboración de conclusiones, etc. El que se diseñen unas u otras depende de que el programa evaluado soporte actividades y procesos complejos de participación. En ese caso podrán realizarse estrategias sofisticadas de estudio de documentos, trabajo en pequeños grupos o por sectores profesionales, elevación de conclusiones a plenarios de jornadas de evaluación del programa, elaboración de documentos provisionales, redacción del informe final, toma de decisiones para la mejora, etc. En este punto, igual que en el resto del proceso de evaluación, los evaluadores deberán tomar sus propias decisiones según las características del contexto en el que desarrollan la evaluación.

En la figura 4 se representa el esquema del proceso de análisis de la evaluación realizada en el CEP de La Laguna, a la que nos hemos referido anteriormente (Gobantes, 2000). De todas las dimensiones de evaluación, en la figura se ejemplifica la referida a *asesoramiento*. En primer lugar se procede a la elaboración de los *principios teóricos* por los que se considera que debe regir el trabajo de asesoramiento por parte del CEP a los centros docentes y al profesorado. En el caso tomado como ejemplo, el proceso de elaboración se realizó con la participación de los asesores y asesoras del CEP puesto que de ello dependía tanto su implicación como el *ajuste entre realidad práctica del asesoramiento y teoría elaborada* (¿a juicio de los asesores y asesoras, la teoría consigue explicar su realidad concreta?).

A partir de ahí, con la referencia ya elaborada, el análisis de la evaluación consiste en contrastar los *principios* con *lo observado* a través de la recolección de información efectuada por la evaluación. Todo ello quiere decir que es fundamental que la evaluación se ocupe, a lo largo del proceso, de elaborar un marco interpreta-

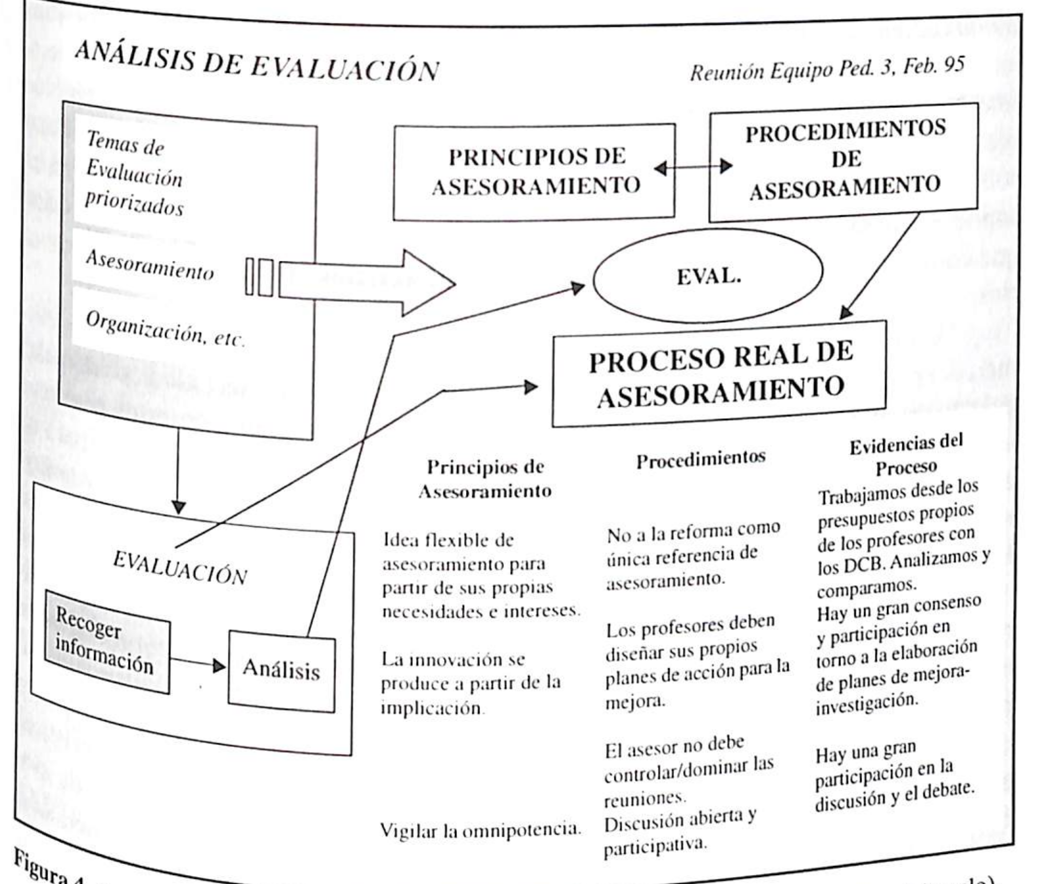


Figura 4. Análisis para la evaluación: relación entre principios-procedimientos-proceso (ejemplo).

tivo (compartido o no con las audiencias según el modelo de evaluación adoptado), en este caso acerca de las características del asesoramiento, a partir del realizar el análisis de los datos recopilados y establecer los puntos en los que el programa debe mejorar. Por lo tanto, la rigurosidad del análisis esta determinada por la recuperación de los criterios y su contrastación sistemática con los datos y evidencias obtenidos.

ACTIVIDADES PARA EL ALUMNO

Teniendo en cuenta los objetivos del capítulo y su contenido, más que proponer varias actividades cortas y puntuales sobre el contenido del capítulo, vamos a proponer una actividad de carácter práctico, que inicie a los alumnos y alumnas en el diseño de evaluación de programas. La actividad consiste en elaborar un diseño de evaluación para un programa previamente seleccionado por el estudiante.

Para ello deberán comenzar por revisar bibliografía y revistas especializadas en las que se describan programas ya realizados o por realizar, de intervención

social (educativos o no). El siguiente paso será elegir uno, mejor si se ajusta a sus intereses intelectuales acerca del tema. Lo que nos interesa no es un tema concreto, esa elección será a juicio del alumno o alumna, sino un programa cuya descripción sea sencilla, clara y contenga los elementos suficientes de su diseño para trabajar sobre él. La finalidad de esta actividad es centrar la elaboración del estudiante en un supuesto concreto para que el trabajo no sea simplemente teórico y abstracto sino que conjugue las propuestas prácticas con las referencias teóricas que sean necesarias.

Una vez hecha la elección, se procederá a la elaboración de un pequeño resumen del programa elegido a modo de presentación de las características principales del objeto que será sometido a evaluación. Seguidamente redactaremos un breve marco teórico acerca de la evaluación de programas. El criterio a seguir para su elaboración consiste en realizar un trabajo personal en el que se *sinteticen creativamente*, en una propuesta teórica, los distintos modelos y prácticas evaluativas. Puesto que la finalidad del marco teórico es argumentar creativamente acerca de la evaluación más idónea para el programa seleccionado, se debe evitar hacer un resumen de cada una de las propuestas de los distintos modelos y autores (como si de un diccionario se tratara). Un trabajo así realizado no aportaría claridad acerca de las intenciones que tenemos con respecto a la evaluación del programa que hemos elegido. Esto último quiere decir que debemos tener en cuenta las características del programa seleccionado para argumentar teóricamente a favor del tipo de evaluación que mejor se ajustará a sus necesidades y peculiaridades. Es recomendable que al final de este apartado se elabore algún tipo de conclusiones acerca de la propuesta teórica de evaluación.

Una vez realizado el marco teórico pasaríamos a desarrollar el esquema, en sus distintas fases (A, B y C), cuyo contenido ha sido explicado en el capítulo. Sugerimos que el esquema se utilice creativamente, es decir, que se introduzcan nuevos elementos, se eliminen los que se consideren o que se modifique el orden que hemos propuesto si se considera necesario y justificado. Eso sí, se trata de trabajar a partir de un esquema previo que se ajuste a nuestras necesidades e intenciones con respecto a la evaluación del programa concreto, por lo que los cambios no deben ser arbitrarios y deberán mejorar, no empeorar, la estructura que hemos propuesto.

BIBLIOGRAFÍA COMENTADA

MIGUEL, MARIO DE (1995). *La calidad de la educación y las variables de proceso y de producto*. Donostia: *Eusko Ikaskuntza, Cuadernos de Sección. Educación* 8. p. 29-51.

Como en el capítulo se indica considero necesario relacionar evaluación con calidad. El trabajo de De Miguel nos introduce en el concepto, clasificando sus

diferentes usos, identificando variables para su evaluación, elaborando un marco teórico de referencia para cada una de las variables identificadas, extrayendo criterios y relacionando las distintas concepciones acerca de la calidad con diferentes modelos de evaluación. El texto que comentamos es de gran ayuda para la elaboración del marco de referencia que sugerimos en el apartado 3.1 de la fase A y para identificar dimensiones, subdimensiones e indicadores de evaluación (2.1 y 3er. punto de la fase C).

COLÁS BRAVO, Mª.P. y REBOLLO CATALÁN, Mª.A. (1993). *Evaluación de Programas. Una guía práctica*. Sevilla: Kronos.

Como se indica en su subtítulo, es una guía práctica que nos ayudará a situarnos de forma sencilla en el entramado de modelos y perspectivas de evaluación. Igualmente es útil como apoyo para la elaboración de propuestas de evaluación como la que se sugiere en la actividad para los alumnos: función de la evaluación, modelos, contenido de la evaluación, diseños, metodología, recogida de información e interpretación, elaboración del informe de evaluación, etc. En los anexos se añaden tres ejemplos de evaluaciones realizadas desde perspectivas muy distintas. Es muy útil para el aprendizaje que el lector se detenga a analizar las diferencias entre ellas acerca de la finalidad, el enfoque, la metodología, el papel de los evaluadores, etc.

ALVIRA MARTÍN, F. (1991). *Metodología de la evaluación de programas*. Madrid: Centro de Investigaciones Sociológicas.

Un texto breve muy bien sistematizado que contribuye a clarificar el modo en el que deben hacerse las evaluaciones atendiendo a la dificultad práctica y complejidad de los programas. Resume diferentes autores clásicos, desarrolla diferentes fases de la evaluación e incluye un esquema para la participación de las distintas audiencias y la utilización de resultados en la toma de decisiones. Desarrolla ampliamente la evaluación de la evaluabilidad, de la implementación, de los resultados para, finalmente, referirse a la monitorización y seguimiento de programas.

Para la ampliación de estos dos últimos libros recomendamos los citados en el texto, tales como, por ejemplo, STUFFLEBEAM y SHINKFIELD (1987), BRINKERHOFF *et al.* (1985), WORTEN y SANDERS (1987).

CAPÍTULO IV

EVALUACIÓN Y CALIDAD DEL PROFESORADO

José Manuel García Ramos[1]
Elvira Congosto Luna[2]
Universidad Complutense de Madrid

OBJETIVOS

1. Reflexionar sobre el concepto de evaluación y sus perspectivas actuales aplicadas a la labor del profesor.
2. Identificar diferentes modelos de evaluación del profesorado, en función de la fundamentación teórica que los sustentan y de sus características aplicadas.
3. Conocer las funciones y usos de la evaluación del profesor en el contexto institucional.
4. Diferenciar los agentes intervinientes en la evaluación docente.
5. Analizar las características diferenciales en los procesos de evaluación en función del nivel o etapa educativa de que se trate, con especial detenimiento en el caso universitario.
6. Estudiar los requisitos que deben contemplar las técnicas e instrumentos para la evaluación del profesor.

CONTENIDOS

---

[1] Facultad de Educación de la Universidad Complutense de Madrid. Área de Métodos de Investigación y Diagnóstico en Educación. Royo Villanueva s/n. 28040 MADRID.
[2] Coordinadora del Departamento de Calidad, Evaluación Institucional y Formación del Profesorado del Centro Universitario Francisco de Vitoria. MADRID.

INTRODUCCIÓN

La evaluación del profesorado constituye uno de los tópicos de mayor interés y actualidad en nuestro país, en el intento de dar respuesta a la necesidad de mejorar la calidad de nuestros centros y a las exigencias del marco legislativo. El análisis de los estudios realizados en este campo, nos lo muestra como uno de los temas con más tradición en la investigación evaluativa y al que se han dedicado más esfuerzos. No por ello se encuentra libre de conflictos, sobre todo cuando el enfoque de la acción se proyecta para la toma de decisiones.

La evaluación del profesorado, de la actividad docente, no debe contemplarse de forma aislada. El docente forma parte de una comunidad educativa en la que desarrolla su labor, viéndose influido por numerosos factores. La mejora de la calidad de una institución educativa no se logra mediante la optimización de forma autónoma de los diversos elementos que la componen, sino mediante la búsqueda y consecución de coherencia entre los múltiples integrantes del sistema educativo, en concreto de sus rasgos generales. Logrando el óptimo grado de coherencia entre los componentes fundamentales del sistema, obtendremos la excelencia educativa.

Un centro de enseñanza, si quiere responder al principio de excelencia académica, debe abordar la institucionalización de un procedimiento valorativo de su actividad, cuya aplicación aporte información acerca de la correlación existente entre los medios utilizados y los resultados obtenidos, información que apoye las decisiones oportunas con el fin de mejorar los aspectos críticos detectados y de potenciar aquellos cuya valoración ha sido positiva. Este proceso no es otro que la evaluación, cuya influencia es decisiva, pues, afecta directa e indirectamente a la calidad educativa.

En todo proceso educativo, y siguiendo a Pérez Juste (1988), han de tomar "cuerpo no sólo esas metas educativas, armónica y ordenadamente articuladas, referidas a la totalidad del hombre, sino, como es natural, una adecuada selección de medios -humanos, técnicos y materiales-, una eficaz organización de las instituciones y una rigurosa evaluación de los planteamientos, contexto, procesos y resultados". La evaluación educativa "debe abarcar todo el conjunto de factores significativos que inciden, no sólo en el aprendizaje, sino en la educación integral de los sujetos. Pero a la vez, debe referirse también a todo el conjunto de elementos que forman parte del proceso educativo, condicionando o facilitando la mejora de los educandos" (García Ramos, 1989). Concebimos, pues, la evaluación del profesorado como parte de la evaluación de la institución educativa (enfoque holístico), para la mejora de la calidad de la formación que ofrece.

Los objetivos que perseguimos en este capítulo son varios. Pretendemos reflexionar sobre el concepto de evaluación del profesorado y sus perspectivas actuales, analizando los distintos modelos que existen, sus funciones y usos. El profesor debe ser el primer interesado en participar en la evaluación de su tarea, ya que el

objetivo fundamental es su mejora. Pero, para obtener una visión más completa es importante conocer la opinión de otros agentes como sus compañeros, también profesores, de sus alumnos e, incluso, de la dirección o administración de la propia institución.

Para lograr involucrar a los profesores en un proceso de evaluación de su labor, enmarcado dentro del plan general de evaluación de un centro, la dirección debe crear y fomentar una "cultura evaluativa" cuyo objetivo final es la optimización de los servicios educativos que se prestan. Por este motivo, debe definir unas normas claras, objetivas y fiables que rijan la recogida de información, el tratamiento de los datos, la elaboración de informes con los resultados obtenidos, la transmisión de éstos, en primer lugar al interesado y, en segundo lugar, a los distintos miembros de la dirección, y, por último, el diseño del plan de formación a nivel personal con el propio docente y a nivel de grupo con los profesores y el equipo directivo. Éste es nuestro segundo y ambicioso objetivo, analizar las características que debe tener el proceso de evaluación del profesor, proceso vivo, abierto a la participación e incorporación de ideas proporcionadas por los interesados.

Finalizaremos analizando cómo se desarrolla la evaluación del profesor en los centros en función del nivel educativo en el que desarrollen su labor, teniendo en cuenta cómo contempla esta actividad la normativa vigente, y las diferentes técnicas e instrumentos que facilitan la recogida de información.

1. CONCEPTO Y PERSPECTIVAS ACTUALES SOBRE LA EVALUACIÓN DEL PROFESORADO

La evaluación del profesorado es, todavía hoy, una realidad de escasa práctica en las instituciones educativas, aunque hemos de reconocer que se trata de una actividad que cada día se va extendiendo más entre los docentes, sobre todo, con el objetivo de mejorar su práctica.

Entendemos la evaluación del profesorado como un proceso y no como una mera actividad puntual. Siguiendo a De la Orden (1992), la evaluación docente sería un proceso sistemático de recogida, análisis e interpretación de información relevante de la labor docente y la formulación de un juicio de valor sobre su adecuación a un criterio o patrón, como base para la toma de decisiones acerca de tal faceta.

La evaluación del profesorado ha de proporcionar un conocimiento de la situación de partida, y del grado de coherencia de sus elementos integrantes. Tras su valoración, ha de elaborarse un programa de mejora de los aspectos susceptibles de la misma. Una tarea como la evaluadora, exige una cuidadosa planificación que vaya desde la especificación de los juicios que deben emitirse a las decisiones que deberán asumirse y darse a conocer a los interesados. Entre el principio y el final del proceso deberán cubrirse una serie de etapas, tales como: recogida de informa-

ción previa, definición de objetivos educativos, determinación de los sistemas de referencia a utilizar para cada objetivo, construcción y selección del procedimiento o técnicas para registrar o medir variables relevantes, aplicación de pruebas, etc.

La evaluación del profesor es, pues, un proceso sistemático y permanente, integrado en la actividad educativa, que ayuda a comprender la realidad, invitando al docente a reflexionar valorativamente sobre su labor y tomar las decisiones oportunas para su perfeccionamiento profesional y, consecuentemente, para la mejora institucional.

Hasta hace poco tiempo, la evaluación del profesorado no era considerada una actividad importante, no se entendía como aspecto vital de la calidad de la enseñanza. En la legislación educativa, la Ley General de Educación (LGE, 1970), en su artículo 11.5, empleaba el término evaluación de centros, aunque no contemplaba de manera directa la evaluación del profesor. Este aspecto cambia con la Ley Orgánica General del Sistema Educativo (LOGSE, 1990) en la que se recoge por primera vez de forma explícita la necesidad de evaluar la tarea del profesor, dentro de un sistema de evaluación del centro.

> *"La evaluación del sistema educativo se orientará a su permanente adecuación a las demandas sociales y a las necesidades educativas y se aplicará sobre los alumnos, el profesorado, el centro, los procesos educativos y sobre la propia Administración"* (art. 62.1).

En el ámbito universitario, la Ley de Reforma Universitaria (LRU, 1983) contempla la evaluación del profesor, pero se centra, a diferencia de la LOGSE, en la individualidad del profesor como objeto de evaluación y no dentro del contexto de una evaluación institucional o de centro.

Actualmente, resultado del interés por la mejora de la "calidad" de los servicios que las empresas ofrecen, entre las que se encuentran, por supuesto, las instituciones educativas, se ha producido un gran desarrollo de la evaluación institucional y, por tanto, de la evaluación de todos sus componentes. Un desarrollo que hace referencia no sólo al incremento del número de centros que incorporan este proceso, sino a la revisión constante del funcionamiento de los sistemas evaluativos diseñados y aplicados, de los distintos métodos propuestos para mejorarlos, y a la creación también de institutos de evaluación que velan por todos estos aspectos en el ámbito educativo.

Centrándonos en la evaluación del profesor, amplio es el número de investigaciones que se plantean cuáles son los aspectos que deben evaluarse. Así encontramos estudios centrados en las aptitudes o características personales, *en cualidades personales que el docente posee y manifiesta*; por ejemplo: facilidad de comunicación y empatía, interés y perseverancia en su trabajo, etc., para que desarrolle su tarea de forma eficaz. Son los denominados *indicadores de presagio (a priori)*.

No se trata de conductas específicas que se producen en el aula, sino de características del comportamiento general del profesor. La investigación parece mostrar que estos indicadores de presagio, que tienen en su fondo suficiente potencial como criterios explicativos o indicadores de eficacia docente, tienen cierta influencia directa sobre las conductas del profesor en el aula y sobre el aula, aunque su influencia es sobre todo indirecta; es decir, se traducen en modos de acción docente, en formas de llevar a efecto determinadas conductas. Las críticas más fuertes que han recibido este tipo de variables como criterios de evaluación hacen referencia a su ambigüedad y mala definición, a que las variables de personalidad tienen escasa influencia por ser muy generales y estar demasiado alejadas de otras dimensiones generales influyentes y, sobre todo, a que estas características se encuentran poco próximas al producto educativo.

Pero no podemos centrar la evaluación del profesor en sus características personales solamente. Debemos abordar otros factores que muestren si está preparado profesionalmente y sabe desarrollar sus *competencias docentes* en el aula y en las restantes tareas que le corresponden teniendo en cuenta el centro educativo al que pertenece y el nivel de enseñanza que imparte.

Quizá la evaluación de la actuación de un profesor en el aula sea el método más común en la actualidad, para evaluar al docente, ya que permite verificar el grado en que las cualidades, actitudes y habilidades didácticas que manifiesta afectan a los resultados de los estudiantes. Pero, antes que la actuación del profesor en el aula, en la evaluación hay que tener en cuenta el dominio de contenidos de la materia que enseña y la programación y preparación de las clases. En este sentido, puede decirse que el "dominio de los contenidos" de su asignatura o disciplina, es la *condición necesaria*, pero la *condición suficiente* para que produzca una acción educativo-formadora eficaz, está ligada al resto de procesos que facilitan la *comunicación* de los contenidos y la *motivación* de aprendizaje que generan en el estudiante. Ambas condiciones son básicas y la segunda presupone a la primera.

La mayor parte de la investigación evaluativa sobre eficacia docente se ha centrado en los denominados *indicadores de proceso*. Los procedimientos seguidos para definir dichos criterios han sido tanto de tipo deductivo (partir de un concepto general y descomponer en características o conductas que podrían servir como criterios) como inductivo (búsqueda de comportamientos docentes en la observación de la clase, esto es, encontrar criterios mediante observación).

Ambos procedimientos han mostrado sus ventajas y sus inconvenientes. Por ejemplo, si se parte de destrezas o habilidades del profesor dividiendo la conducta en otras más pequeñas, tenemos la ventaja de que hay una baja inferencia en los resultados, esto es, si las correlaciones de esas microconductas con el logro son altas, puede establecerse con precisión dicha relación; el problema suele estar en que dichas microconductas no suelen ser representativas de la conducta global, son elementos poco significativos. Por otra parte, puede llegarse a una gran atomiza-

ción de las variables que resulta poco operativa e inmanejable en cualquiera de sus propósitos con que se quisieran utilizar. Campbell y Barnes (1972) han llegado a contabilizar hasta un total de seiscientos microelementos, aislados en el conjunto de los distintos esquemas de observación.

Toda la labor del profesor repercute en los estudiantes. Sin embargo, éste es un punto discutible y crítico, por la dificultad que conlleva separar las distintas variables o indicadores que influyen en la conducta de los alumnos *(indicadores de producto).*

Éstos son algunos de los criterios más estudiados y utilizados en los procesos de evaluación del profesorado. Pero, la lista puede ampliarse y debe incluir aquellos que la propia institución considere relevantes en su equipo. Lo que nos ha de quedar claro, es la necesidad de emplear diversos criterios y fuentes, de tal forma que el conocimiento que tengamos de la labor del docente sea lo más amplia y objetiva posible. Para comenzar, cada institución, basándose en la literatura sobre el tema y en su propio concepto de profesor eficaz (que surge de su naturaleza como centro educativo, de su proyecto educativo y su misión, de sus características propias), debe elaborar su propuesta teórica y justificada de núcleos a valorar y de criterios e indicadores a incluir en su propuesta evaluadora.

2. MODELOS DE EVALUACIÓN DEL PROFESORADO

Los sistemas de evaluación del profesorado que se desarrollan en los centros educativos responden a diferentes modelos, ni mejores ni peores, que centran más la atención en unos determinados aspectos, de acuerdo con los objetivos definidos en la cultura evaluativa de cada institución.

Escudero Escorza (1991), citando a De Miguel (1989), indica que un verdadero modelo requiere la fundamentación del objeto de evaluación en una teoría sobre la estructura y funcionamiento del objeto de evaluación a través de variables debidamente relacionadas y operacionalizadas, con expresión de validez empírica, criterios de eficacia y campo de aplicación.

Son varios los modelos de evaluación docente existentes, en función de la fundamentación teórica que lo sustenta y de las características de su aplicación a la práctica. Así Marczely (1992), citado por Mateo y otros (1996), detecta siete modelos de evaluación del profesor:

1. El centrado en rasgos del profesor.
2. El de escala de valoración (comprueba el grado en que un rasgo o conducta están presentes).
3. El centrado en los objetivos de rendimiento.
4. El centrado en los objetivos instruccionales.
5. El orientado hacia el proceso.

6. El centrado en las preocupaciones del profesor (dando oportunidad a que los profesores analicen su propio progreso como profesionales y busquen apoyos para la mejora).
7. El ecléctico (combinación de dos o más enfoques).

Cada modelo ha surgido en un contexto definido y parte de una concepción educativa particular. Cada modelo responde de una manera particular a un enfoque de la evaluación institucional y de la evaluación docente en particular. En cada caso, y desde una perspectiva aplicada, los responsables de definir y poner en marcha dichos procesos deberán elaborar su propio modelo acorde con sus planteamientos y con las funciones y objetivos que se planteen.

No existe, pues, un modelo de evaluación único y generalizable, dada, incluso, la continua evolución del concepto mismo de evaluación y la gran cantidad de modelos distintos de evaluación que han surgido durante los últimos treinta años.

El modelo de Tyler (1943) cuyos elementos esenciales son tres: objetivos (o conductas deseables), actividades de aprendizaje (individuales o de grupo) y la calificación de los aprendizajes (rendimiento), sirvió de base para la formulación de nuevos modelos. Modelos como los de Provus (1971), Hammond (1972), desarrollo de las taxonomías (Bloom, 1956), objetivos instructivos de Popham (1969) y el modelo multietápico de Mettfesel y Michael (1967), fueron derivaciones del modelo de Tyler. No obstante, las cuestiones en torno a las relaciones entre "cumplimiento" y "no cumplimiento" de objetivos y "eficacia" versus "ineficacia" de los programas, departamentos o servicios, llevó a modificar el original modelo de Tyler, considerando la necesidad de evaluar los objetivos mismos de los programas (correctos/incorrectos; utópicos/factibles; relevantes/irrelevantes), así como la inclusión del concepto de necesidades en el propio modelo.

Surgieron también los modelos orientados al juicio, basados en la competencia valorativa y profesional del evaluador para emitir juicios. Sin embargo, la realidad hizo ver que el mundo de la programación social y política, dentro del que está inmersa la educación, se basaba en criterios valorativos previamente seleccionados y justificados, por lo que la pretendida "objetividad" desacreditó los estudios evaluativos (Carballo, 1990).

En esa época Scriven defendió el concepto de valoración, como característica ineludible de toda evaluación. Surgieron otras alternativas (Stake, 1967; Eisner, 1971) que asignaban al evaluador el papel de "crítico educativo" y algunos modelos políticos o de debate (Owens, 1973; Wolf, 1974) que tuvieron gran repercusión en estudios tan conocidos como el "Informe Coleman" (Coleman y otros, 1966) o el "Segundo Informe Coleman" (Jenks y otros, 1972), de gran trascendencia por la valoración crítica de la eficacia de las instituciones educativas norteamericanas (Fuentes, 1986); no obstante, estos últimos modelos aparecían "viciados" por el ambiente político en que se desarrollaban. También aparecieron los modelos de evaluación de costos, cuya finalidad era determinar las necesidades de inversión

económica en cada programa y así elegir el más barato. El problema resultó ser la fuerte relación entre calidad y coste.

En los años ochenta, el énfasis se puso en la evaluación como soporte de la toma de decisiones administrativas, políticas y, por supuesto, educativas. Junto a esta corriente, otra puso su centro de atención en el compromiso de la evaluación y del evaluador, con la mejora de los programas y las instituciones (Stufflebean y Shinkfield, 1987; Carballo, 1990).

La primera posición reforzó modelos tan importantes como la planificación educativa a través de "UTOS" (Cronbach, 1963 y 1982); Modelo CIPP (Stufflebean, 1967 y 1985); Modelo centrado en el cliente (Scriven, 1967); Modelo de evaluación formativa (Brinkerhoff y otros, 1983); y otros estudios de numerosos autores como Alkin (1969), Reichard (1972), Taylord (1984), Guba (1978), Webster (1975), etc. Esta postura fomenta la utilización de la evaluación de una forma continua y sistemática, poniendo a disposición de los usuarios los servicios que satisfagan sus necesidades. Asimismo, parece poner el énfasis fundamental en la función sumativa de la evaluación, ayudando al tomador de decisiones a valorar el programa, destacando los puntos fuertes y débiles, una vez el programa ha finalizado.

La segunda postura reivindicó más el compromiso de la evaluación con la mejora de los programas e instituciones, promoviendo modelos en los que el punto de vista del evaluado era esencial (Evaluación Pluralista o Iluminativa, de Parlett y Hamilton, 1977; Evaluación Respondente de Stake, 1975; Evaluación Transaccional de Rippey, 1973; y Evaluación Democrática de McDonnall, 1975). Esta posición conduce a investigaciones activas, participativas y de gran riqueza informativa, tendiendo a proporcionar a los departamentos o servicios evaluados, los recursos necesarios para dirigir sus propias evaluaciones. Parece, por tanto, poner el énfasis en la mejora, en el cambio, función formativa de la evaluación, al ofrecer su apoyo y ayuda a las personas involucradas en el centro o en el programa, para su perfeccionamiento, mientras se desarrolla la acción de los mismos.

En todo caso, hoy aceptamos que ambas funciones, sumativa y formativa son necesarias y complementarias. Ambas deben estar presentes en un proceso de evaluación institucional.

3. FUNCIONES Y USOS DE LA EVALUACIÓN

Una revisión de las diferentes funciones que se atribuyen a la evaluación del profesor en la literatura sobre el tema nos lleva a comprobar la gran variedad existente. Por ejemplo, Scriven (1967) fue el primero que sugirió la diferenciación entre *evaluación formativa* y *sumativa*, distinguiendo las dos principales funciones que debe cumplir la evaluación si tenemos en cuenta su finalidad. En la evaluación *formativa*, se recogen de forma rigurosa datos sobre la labor realizada por el profesor, con la finalidad de tomar decisiones para su mejora, mientras el programa o

proceso de formación del profesorado está en marcha. En esta línea, la evaluación formativa es parte de un proceso más general de formación del profesorado, definido al nivel de la institución. Se trata de un tipo de evaluación de carácter continuo.

La diferencia con la evaluación *sumativa* es que en ésta no se contempla la posibilidad de modificar para mejorar el proceso evaluado, sino para mejorar procesos futuros. Algunos autores señalan a este respecto que, cuando la evaluación viene "desde las altas instancias" es vivida por el equipo docente como un "rendimiento de cuentas" (McLaughlin y Pfeifer, 1988). En ocasiones, esta función de la evaluación es incorporada a las políticas institucionales y departamentales para considerar aspectos como la promoción, la titularidad, los ajustes salariales, etc.

Dentro del ámbito de la evaluación que proporciona información para la mejora, el profesor debe reflexionar sobre su práctica docente como primer paso para su perfeccionamiento profesional y, consecuentemente, para el incremento de la calidad de la docencia en general. Esta reflexión se convierte en una herramienta de diagnóstico de los puntos débiles y fuertes del profesor (*función diagnóstica* de la evaluación). Para alcanzar este objetivo, hemos de conseguir en el docente una fuerte motivación interna para el aprendizaje, como actitud abierta al cambio, a la búsqueda de soluciones a las dificultades y al diseño de estrategias concretas para favorecer el perfeccionamiento docente en destrezas y habilidades instructivas y formativas. Si se logra extender esta cultura de evaluación entre los profesores, se conseguirá mejorar la institución educativa en la medida en que las metas y objetivos personales se integren con los institucionales. Como señalan Villa y Morales (1993):

> *"A nivel institucional, la evaluación del profesorado estimula la acción porque cada nueva evaluación proporciona una nueva oportunidad para aprender y para definir criterios de actuación aceptables dentro de la institución o conjunto de instituciones".*

Otro criterio clasificatorio de los tipos de evaluación existentes, independientemente del modelo de evaluación docente que sigamos, es el "cuándo" se realiza la evaluación. Así tendremos como modalidades básicas, *inicial*, *continua* o *progresiva* y *final*. Es evidente que, en el fondo, las tres modalidades podrían entenderse también como pasos de un mismo proceso evaluativo global, cuyo referente básico sería la evaluación progresiva o continua. En esta perspectiva, la evaluación inicial sería considerada como el primer paso de la evaluación progresiva, siendo la evaluación final, desde un punto de vista pedagógico, una resultante de la evaluación continua. Se identifica prácticamente la evaluación progresiva con la formativa, mientras que la evaluación sumativa suele ordenarse a la toma de decisiones al final del proceso, por lo que se encuentra más vinculada a la evaluación final. En este sentido, evaluación sumativa y formativa pueden ser complementarias, aunque

todo sistema de evaluación institucional debe poner su acento en una de ellas, sin descartar el uso de la otra función.

Una tercera clasificación, vinculada a las anteriores de alguna manera, distingue entre evaluación *normativa*, *criterial* y *personalizada*, siendo las dos primeras modalidades de *heterorreferencia*, y la última, con referencia a las características y posibilidades personales (*autorreferencia*).

Estas clasificaciones anteriores no agotan, sin embargo, las modalidades de evaluación del profesorado que pueden encontrarse en los manuales sobre la materia. Algunas otras dicotomías que suelen mencionarse son: cuantitativa-cualitativa, interna-externa, heteroevaluación-autoevaluación y orientadora-prescriptiva.

El enfoque *cuantitativo* hace referencia a aquellos procedimientos que requieren de la medición y cuantificación de los diferentes aspectos a tener en cuenta en la evaluación del profesorado. Intentan identificar las relaciones causa-efecto entre las variables dependientes e independientes, consideradas susceptibles de generalización, aspecto este último muy discutido por aquellos especialistas más partidarios de observar el fenómeno en su complejidad, teniendo en cuenta las características y circunstancias peculiares de cada situación. Se trataría en este caso de un enfoque *cualitativo*. Ambos enfoques, sin embargo, pueden ser complementarios en un sistema particular.

También podemos distinguir entre *evaluación interna*, a cargo de los mismos agentes implicados en el proceso evaluativo, y la que se efectúa desde fuera, a través de jueces externos, *evaluación externa*. La evaluación interna conlleva una implicación mayor de los miembros de la comunidad educativa por lo que permite al docente, y por ende a la institución, reflexionar sobre su labor, impulsar el diálogo y la participación. Normalmente, y a diferencia de la externa, no es vista como una amenaza, una intromisión o un juicio. Sin embargo, el profesorado adolece generalmente de formación técnica para evaluar su labor, dificultad que normalmente no aparece en la evaluación externa porque es realizada por expertos en la materia. Existe la posibilidad de evitar este aspecto solicitando a iniciativa de la propia institución, una evaluación externa que suele requerir la colaboración y compromiso de los profesores del centro porque también es demandada por ellos. La finalidad de la misma sería eminentemente formativa. Como vemos, ninguna de ellas tiene todas las ventajas o todas las desventajas. La elección, por tanto, deberá realizarla la institución educativa, sopesando los diversos factores que inciden en cada caso. Algunos de estos factores, siguiendo a Weis (1975) son: la objetividad, tranquilidad o confianza de los administradores, comprensión del programa, potencial de utilización y autonomía (independencia de juicios). Es evidente que, dependiendo del grado de importancia que se den a estos factores, se decidirá uno u otro modelo.

Si tenemos en cuenta el agente que realiza la evaluación nos encontramos con la diferenciación entre *heteroevaluación* y *autoevaluación*. En la primera, los

agentes evaluadores más comunes son la dirección (más frecuentemente la inspección), los compañeros del equipo docente de ciclo o de departamento, e incluso el claustro. A diferencia de la heteroevaluación, en la autoevaluación es el propio profesor quien reflexiona sobre su quehacer docente y se evalúa a sí mismo. Dedicaremos más adelante un apartado para ver las características de la autoevaluación docente.

Por último, cabe también establecer la distinción entre aquellas evaluaciones cuya meta es esencialmente *orientadora* (formativa) y otras más bien *descriptivas* (de naturaleza sumativa fundamentalmente), dirigidas a proporcionar una conclusión precisa sobre determinados aspectos de la labor docente.

Independientemente del modelo de evaluación docente que sigamos, el profesor ha de sentirse partícipe del proceso evaluativo: debe conocer los propósitos de la evaluación y tener información sobre el método y los procedimientos que se van a emplear; debe poder realizar aportaciones en el desarrollo del sistema de evaluación. El plan de evaluación del docente debe ser riguroso, sistemático y estructurado, desarrollado por expertos y personal técnico y proporcionarle los servicios de apoyo necesarios y los medios que le lleven al perfeccionamiento de su labor. De este modo, el docente se comprometerá con la institución educativa y sentirá el apoyo del equipo directivo. El modelo de evaluación debe permitir un diálogo abierto dentro de la organización. En este sentido, conviene no separar en una institución las funciones evaluadoras de las de formación del profesorado. Ambas funciones contribuyen decisivamente a la calidad del profesorado y de la institución y deben ser responsabilidad del mismo departamento de la institución.

4. AGENTES DE LA EVALUACIÓN DEL PROFESOR

En la bibliografía sobre la materia, se definen numerosas opciones en cuanto a quiénes son los agentes de la evaluación del profesorado. En este apartado nos referiremos a las principales que son los estudiantes, los colegas y el propio profesor.

Podemos indicar otras como los administradores y directivos académicos o los departamentos. En este caso consideramos importante señalar que el departamento no debe acometer la evaluación individualizada de cada profesor, sino limitarse a acotar los ámbitos en donde la recomendación de cambios está asociada a la mejora de la calidad del trabajo del departamento (Rodríguez Espinar, 1991).

4.1. La evaluación realizada por los alumnos

Las evaluaciones de los alumnos constituyen la fuente más frecuentemente utilizada para la evaluación de la calidad docente, sobre todo en la universidad. Ramsden (1991), desde la perspectiva del aprendizaje del alumno, señala que éste es el único capaz de valorar si la docencia recibida le ha facilitado el aprendizaje.

Considera que el alumno es capaz de distinguir entre buena actuación del profesor y docencia efectiva.

Sin embargo, somos conscientes de que este tipo de evaluación no siempre es bien acogida por aquellos profesores que se preocupan más por los factores que puedan influir en los resultados de la evaluación (subjetividad de los alumnos), que en analizar con detenimiento los resultados obtenidos en dichas valoraciones. Los estudios de validación comparan estas evaluaciones con otros criterios de calidad o con otras características que pueden influir indebidamente en la evaluación. A pesar de las críticas, toda la investigación hecha sobre estas valoraciones de estudiantes indica que son razonablemente válidas y que no se ven contaminadas en un grado apreciable por factores distintos a la misma calidad docente del profesor (Villa y Morales, 1993). Incluso, puede considerarse la vía más fiable y objetiva de las posibles.

Por lo que respecta a la fiabilidad, Kurz et al. (1989) señalan que el grado de unanimidad de los alumnos al diferenciar unos profesores de otros, presenta un elevado nivel en cuanto a la consistencia interna de las escalas, entre sujetos, diferentes momentos del curso, y diferentes cursos de la misma tipología. Fernández Díaz (1988) recoge las conclusiones de los estudios de Murray (1984), quien indica que las valoraciones de la enseñanza de la clase correlacionan de forma moderada a alta (de 0,50 a 0,90) con valoraciones comparables hechas por supervisores, compañeros, alumnos y observadores externos de clase indicando que las percepciones "buena" y "pobre" enseñanza son similares a las de los observadores más expertos, de mayor madurez y neutralidad. También se ha comprobado que las valoraciones de los estudiántes muestran correlaciones pequeñas pero significativas con factores tales como el tamaño de la clase, nivel de curso, etc. Aunque los resultados son algunas veces inconsistentes, el peso de la evidencia indica, al menos, una correlación positiva moderada entre las valoraciones del estudiante sobre la enseñanza recibida y las medidas objetivas del rendimiento del estudiante.

De acuerdo a las investigaciones más recientes, se puede concluir que los estudiantes son capaces de identificar las dimensiones más relevantes para la definición de lo que se considera una buena enseñanza (Miller, 1987; Marlin, 1987; Marsh, 1984 y 1987, entre otros). Son los mejores jueces de aspectos relacionados con los métodos de enseñanza; equidad pedagógica; actitud, interés y expectativas sobre el alumnado, etc. Asimismo, sus opiniones sobre la enseñanza recibida se asemejan a las de otros agentes implicados en similares contextos educativos (Nackburn y Clark, 1975; Doyle y Crichton, 1978; Marsh y otros, 1979; citados por Mateo y Fernández, 1992), permanecen relativamente estables con el trascurso del tiempo (Centra, 1974; Marsh, 1977) y parecen correlacionados con el nivel de aprendizaje alcanzado (Centra, 1977; Cohen, 1981; Marsh y Overall, 1980).

Somos conscientes de la existencia de una serie de factores que pueden influir en las valoraciones que realizan los estudiantes como el ya señalado efecto de

las calificaciones obtenidas (una medida para disminuir su efecto consiste en realizar las aplicaciones dejando pasar un tiempo prudencial desde la publicación de las notas), o algunas características de personalidad del mismo profesor (liderazgo, extraversión, entre otros).

En todo caso y como recomendación, aconsejamos el uso habitual junto a esta fuente de otras que deben completar la información y contrastarla. Asimismo, deben combinarse no sólo *fuentes*, sino también otras *técnicas* distintas del cuestionario (Franke Wikberg, 1990): observaciones directas de colegas, análisis de materiales de enseñanza, la percepción por parte del propio profesor de su labor docente, la valoración del profesor respecto a su relación con el grupo, entrevistas con los estudiantes, etc.

4.2. La evaluación a través de los colegas

La evaluación por iguales ha sido una práctica muy extendida para juzgar la calidad en una gran variedad de profesionales. Sin embargo, en el mundo educativo esta situación ha sido excepcional. Muchos autores consideran adecuados a los profesores compañeros para evaluar la calidad del contenido del curso. El hecho de conocer la temática, su familiaridad con el currículum, su experiencia profesional y su propia práctica de la enseñanza son características que podrían contribuir a una evaluación adecuada de la enseñanza (Fernández Díaz, 1988).

La evaluación del profesor por los colegas, siguiendo a Villa y Morales (1993) proporciona una perspectiva diferente de la de los superiores y observadores externos, de las autoridades académicas y de la de los estudiantes. Son buenos analistas de aspectos relacionados con la calidad académica del curso, de los materiales didácticos, de la metodología evaluativa y de la actividad académica del equipo docente. Sin embargo, estas valoraciones presentan problemas de fiabilidad y escaso soporte empírico para su validez.

Algunas de las conclusiones a que han llegado las investigaciones en este campo (Marsh, 1984; French-Lanzovick, 1981) indican que se trata de una evaluación incompleta porque es realizada a partir del conocimiento global que los profesores tienen de sus compañeros; conocimiento general y algo subjetivo, que puede estar influido por información que los propios estudiantes transmiten, y puede verse beneficiada por la "generosidad".

Para evitar estos aspectos, la valoración no ha de basarse en una visita a la clase que imparte el docente objeto de la evaluación, sino que esta actividad (realizada tras cierto entrenamiento y con uso de escalas de observación) debe complementarse con evaluación de documentos y otros elementos.

En general, siendo una fuente complementaria importante, no aconsejamos usarla nunca como fuente única. Barber (1990) propone un híbrido entre la revisión de colegas y la autoevaluación. Con la debida prudencia y aceptación del profeso-

rado, puede añadir matices que favorezcan un clima de mejora común del equipo docente en el conjunto de una institución o departamento.

4.3. Autoevaluación docente

La autoevaluación de la función docente dirigida a la mejora del proceso de enseñanza-aprendizaje y de los factores asociados, es la forma más común de todas las evaluaciones del profesorado y es absolutamente necesaria si se quiere una mejora de la calidad de la enseñanza, ya que la evaluación enfocada a la mejora requiere que el profesor desee participar y cambiar. Tiene una finalidad eminentemente formativa (McColskey & Egelson, 1993; Nieto, 1996) y exige el total convencimiento del profesor para la potenciación de sus puntos fuertes y el cambio y la mejora de los aspectos de su práctica que lo precisen. Exige también de una mayor dedicación del profesor, tiempo del que en muchas ocasiones no dispone, por lo que se puede ver obligado a apoyarse en indicadores fáciles de observar, pero poco válidos. Los propios profesores constituyen una fuente fundamental de información sobre aspectos de contextualización de la enseñanza, planteamientos curriculares, métodos didácticos, etc.

El principal problema que nos podemos encontrar en este tipo de evaluación es la alta valoración del profesor respecto a su labor, como tendencia a verse y a mostrarse a sí mismo como eficiente o muy eficiente. Uno de los motivos a que puede deberse es que en los planes de formación del profesorado raramente se incluye formación y práctica en autoevaluación. Otro, la misma naturaleza de la autoevaluación, que, por definición, no puede arrojar resultados muy negativos, que conlleven cierta conciencia de fracaso del profesor que se autovalora.

A través de este proceso, un profesor de manera individual o con la ayuda de otros colegas, se sensibiliza, reflexiona sobre su propia práctica docente en aquellas áreas que él mismo ha identificado y sobre la adecuación y efectividad de su actuación, con la única finalidad de mejorarse a sí mismo. Es la evaluación *por* y *para* el profesor, en la que se fomenta la comprensión y el crecimiento profesional (Osteman & Kottkamp, 1993). Él mismo es quien recoge, interpreta, pondera los distintos factores que intervienen en la enseñanza, valora la información sobre sus conocimientos, destrezas, habilidades y eficacia, conforme a unos criterios que establece, y diseña su plan de formación para la mejora continua y optimización de su desarrollo personal. Este sistema, puede estimular la motivación para el cambio más que cualquier otro proceso de evaluación, pero implica colaboración técnica de expertos.

Elliot (1986) citado por Yáñiz (1998), distingue tres tipos de autoevaluación con diferentes consecuencias para el desarrollo profesional de los docentes:

a) Autoevaluación basada en el conocimiento práctico táctico.
b) Autoevaluación basada en un conocimiento explícito de reglas técnicas.

c) Autoevaluación como deliberación práctica, fundamentada en una responsabilidad moral.

Villa y Morales (1993) recogen los *cinco enfoques* que pueden utilizarse a la hora de realizar la autoevaluación según Carroll (1981): *cuestionarios* de autoevaluación (muy usados para realizar comparaciones con otras fuentes de datos como las evaluaciones de los alumnos, observaciones en clase, etc.), autoinformes (empleo de preguntas de respuesta libre relacionadas con la enseñanza que requieren respuestas concisas), materiales para el *auto-estudio*, *observación de los colegas* y *reproducción visual y oral*. A este listado podemos añadir, diarios (o registro de acontecimientos de clase u otras actividades), datos sobre el rendimiento del alumnado (calificaciones de exámenes, trabajos, etc.); portfolios del profesorado (se emplean para documentar una extensión de práctica docente o competencia, o documentar el cambio o el perfeccionamiento en su área o destreza concreta; incluiría planes de clase, grabaciones de las sesiones de clases, reflexiones personales del profesor sobre su práctica docente, etc.), etc.

En general, recomendamos este ejercicio de autoevaluación, por lo beneficioso que resulta para el profesor y porque le permite contrastar sus valoraciones con las de otras fuentes.

5. NORMAS EN LA EVALUACIÓN DEL PROFESORADO

Como hemos venido reflexionando en los apartados anteriores, el interés básico del proceso de evaluación del profesorado es doble. Por una parte, pretende obtener información objetiva, fiable y válida de la labor docente del profesor, y, por otro, emplear los resultados con el fin de diseñar estrategias formativas para su perfeccionamiento.

La evaluación del profesor debe quedar investida del quehacer científico tanto en el planteamiento y delimitación del problema como en la planificación del diseño de actuación, en la actuación misma acorde con lo planificado y en la recogida de datos con técnicas fiables y válidas, sin olvidar, claro está, su adecuado tratamiento e interpretación para la toma de decisiones.

Muchos han sido los intentos de elaboración de normas que regulen la evaluación educativa. En 1975 se organizó en Estados Unidos el Joint Committee on Standars for Evaluation, compuesto por organizaciones profesionales relacionadas con el ámbito educativo (asociaciones de profesores, consejos escolares, expertos en evaluación, etc.). Este comité, en 1981, publicó una serie de normas para la evaluación de programas, proyectos y materiales, elaboradas por un grupo de expertos presidido por Stufflebeam. Posteriormente, en 1988, elaboró unas normas para la evaluación de personal.

Siguiendo a Nevo (1997), el Comité propuso 30 estándares, divididos en cuatro categorías cuyo orden no implica grado de importancia:

- *De utilidad*: para asegurar que la evaluación satisface necesidades prácticas de información.
- *De viabilidad*: para asegurar que la evaluación es realista y prudente.
- *De propiedad*: para asegurar que la evaluación se realiza legal y éticamente.
- *De precisión*: para asegurar que la evaluación revela y transmite información técnicamente adecuada.

La evaluación del profesor debe estar al servicio de los principios educativos e institucionales, por lo que su diseño debe realizarse en el contexto de la evaluación institucional. La organización, los procedimientos y los indicadores o criterios que se vayan a emplear en la evaluación deben ser acordados entre los interesados y ser coherentes con los objetivos de la institución y el fin de la evaluación. No hemos de olvidar que se trata de una actividad que ayuda a los centros educativos y a sus profesores a mejorar los servicios que prestan. Los objetivos de la evaluación deben ser definidos explícitamente.

Este proceso evaluador debe estar coordinado y ejecutado por personal cualificado profesionalmente, para que el análisis y seguimiento de los procesos de recogida, análisis e interpretación de la información sean válidos y fiables, evitando cualquier tipo de sesgo. La comunicación de los resultados de la evaluación a los interesados debe realizarse en informes claros y precisos y contar, además, con la posibilidad de que los expertos aclaren cualquier tipo de consulta sobre los mismos y puedan actuar en consecuencia. Ante todo, debe asegurarse la confidencialidad de los datos.

Por último, debe establecerse un doble proceso. Por una parte, el diseño de un programa de mejora en función de los resultados obtenidos y, por otro, el seguimiento de los evaluados para que reciban la ayuda necesaria en función de sus necesidades específicas.

6. LA EVALUACIÓN DEL PROFESORADO DE EDUCACIÓN INFANTIL, PRIMARIA Y SECUNDARIA

Como hemos venido analizando, la evaluación del profesorado debe ser un complemento de la formación que favorezca el desarrollo profesional docente. Gracias a ella, detectamos los puntos fuertes de la actividad del profesor y también aquellos susceptibles de mejora; sus mejoras y sus logros.

Los procesos de evaluación, sin embargo, reúnen una serie de características diferenciales en función del nivel o etapa educativa de que se trate, que vienen marcadas por la propia reglamentación y legislación educativa, que detalla, incluso, los aspectos a evaluar y los agentes que deben participar en el proceso evaluador.

En los primeros niveles de enseñanza no existe tradición de diseñar programas de evaluación del profesorado. Sin embargo, la legislación educativa actual

contempla como una exigencia esta actividad, dentro del contexto de la evaluación del sistema educativo.

Así, en la *Educación Infantil*, la evaluación de la práctica docente deberá contemplar todo aquello que incida en la adecuación del proceso de enseñanza a las características y necesidades educativas del alumnado, que ayude a la adaptación del Proyecto Curricular y de la Programación de Aula a estas necesidades, y, en función de ello, realizar las mejoras pertinentes en la actualización docente. Se reseña, asimismo, que debe tener un carácter continuo y formativo (artículos 10-14 de la O.M. 25.820 de 12 de noviembre de 1992, B.O.E. del 21). Fundamentalmente se evaluará:

a) La organización del aula y el ambiente entre los niños y niñas, y la relación entre el profesorado y el alumnado.
b) La coordinación entre los maestros de un mismo ciclo y la coherencia con otros ciclos.
c) La regularidad y calidad de la relación con otros padres.

En *Educación Primaria* (artículos 15 y 16 de la O.M. 25.821 de 12 de noviembre de 1992 -BOE del 21- sobre evaluación en Educación Primaria), los maestros evaluarán los procesos de enseñanza y su propia práctica docente en relación con el logro de los objetivos educativos del currículo. Igualmente evaluarán el desarrollo del proyecto curricular emprendido. Por tanto, el plan de evaluación de la práctica docente, que incluirá precisiones acerca del momento en que la referida evaluación ha de efectuarse y sobre los instrumentos para llevarla a cabo, será propuesto por la Comisión de Coordinación Pedagógica al Claustro de Profesores para su aprobación. Esta evaluación incluirá, además del análisis específico del proyecto curricular:

a) La organización del aula y aprovechamiento de recursos.
b) El carácter de las relaciones entre alumnos/profesores, entre los mismos maestros y la convivencia entre alumnos.
c) El grado de coordinación entre el equipo directivo, Claustro de Profesores, Comisión de Coordinación Pedagógica, Tutores y Maestros especialistas y de apoyo.
d) La regularidad y calidad de la relación padres/tutores.

En el ámbito de la *Educación Secundaria Obligatoria* los programas de evaluación del profesorado persiguen diferentes propósitos como:

a) Desarrollo profesional del docente.
b) Control del cumplimiento/control administrativo.
c) Incentivación.
d) Carrera docente.
e) Selección y estabilidad del profesorado.
f) Sanciones.

Sin embargo, en la práctica, la mayoría de los sistemas de evaluación del docente se desarrollan con un enfoque eminentemente formativo, para el desarrollo

profesional y personal del docente, con la consecuente mejora de la institución en general. Para conseguir esta finalidad se requiere la implicación de los profesores en el proceso evaluador, precisando, como señalan Stiggins y Duke (1988) citados por Mateo y otros (1996):

a) Fuertes expectativas profesionales.
b) Orientación positiva hacia el riesgo.
c) Disposición al cambio y a la experimentación en clase.
d) Apertura a la crítica.
e) Preparación académica y pedagógica.
f) Experiencias previas positivas con la evaluación del profesorado.

Según la normativa vigente para este nivel educativo (artículo 29 de la O.M. 25.707 de 12 de noviembre de 1992 sobre evaluación de la Educación Secundaria Obligatoria -B.O.E. 20 de noviembre), los principales aspectos a evaluar de los procesos de enseñanza y de la práctica docente son:

a) La organización y aprovechamiento de los recursos del centro.
b) El carácter de las relaciones entre profesores y alumnos y entre los mismos profesores, así como la convivencia entre los alumnos.
c) La coordinación entre los órganos y personas responsables, en el Centro, de la planificación y desarrollo de la práctica docente: Equipo Directivo, Claustro de Profesores, Comisión de Coordinación Pedagógica, Departamentos o Seminarios y Tutores.
d) La regularidad y calidad de la relación con los padres o tutores legales.

Como observamos, no sólo se tienen en cuenta aspectos propios del profesor (como rasgos de personalidad, formación) o de la relación con los alumnos en el aula, sino de los distintos aspectos intervinientes en el contexto del centro en el que desarrollan su labor. Los resultados tanto de la evaluación de los aprendizajes de los alumnos como del proceso de enseñanza deben servir para modificar aquellos aspectos de la práctica docente y del Proyecto Curricular de Etapa poco adecuados a las características de los alumnos y al contexto del centro (artículo trigésimo segundo).

Los principales criterios empleados en la evaluación del profesorado en estos niveles de enseñanza son:

a) Planificación y organización de la enseñanza.
b) Motivación, relación y expectativas hacia los estudiantes.
c) Empleo de herramientas audiovisuales u otros recursos que faciliten el aprendizaje de los alumnos.
d) La integración y colaboración con el equipo de profesorado y otros miembros de la institución educativa.
e) Las relaciones con los padres de familia.

En cuanto a los resultados obtenidos de este sistema de evaluación, éstos deben ser el objetivo fundamental de la reflexión del profesorado, para la mejora

del proceso enseñanza-aprendizaje. Para ello, el docente debe haber contado con la colaboración del evaluador, quien será el encargado, entre otras funciones, de velar por la comunicación de todas las partes interesadas en el proceso, motivarle en su desarrollo profesional y programar conjuntamente acciones de mejora personal. Paralelamente, el equipo directivo debe elaborar un programa de formación que ayude a su equipo docente.

El encargado de supervisar y asesorar este proceso de evaluación es el Servicio de Inspección, en las visitas que realice a los centros escolares. La Inspección Educativa tiene como objetivos fundamentales:

a) Ejercer la función de control de la normativa legal como garantía del derecho de los ciudadanos a la educación. La finalidad de este control es que el Sistema Educativo cumpla los objetivos encomendados por la propia sociedad.
b) Realizar la evaluación del funcionamiento y de los resultados de centros educativos, así como de los profesores.
c) Asesorar y orientar para la aplicación de las normas que regulan la actividad escolar a fin de lograr la mejora de la calidad de la enseñanza.

El artículo 36 de la Ley Orgánica 9/1995, de 20 de noviembre, de la participación, la evaluación y el gobierno de los centros docentes señala, entre las funciones de la inspección educativa:

a) Colaborar y supervisar, desde el punto de vista pedagógico y organizativo, el funcionamiento de los centros educativos, tanto de titularidad pública como privada.
b) Colaborar en la mejora de la práctica docente y del funcionamiento de los centros, así como en los procesos de reforma educativa y de renovación pedagógica.
c) Participar en la evaluación del sistema educativo, especialmente en la que corresponde a los centros escolares, a la función directiva y a la función docente, a través del análisis de la organización, funcionamiento y resultados de los mismos.
d) Velar por el cumplimiento, en los centros educativos, de las leyes, reglamentos y demás disposiciones vigentes que afecten al sistema educativo.
e) Informar sobre los programas y actividades de carácter educativo promovidos o autorizados por las administraciones educativas competentes así como sobre cualquier aspecto relacionado con la enseñanza que le sea requerido por la autoridad competente o que conozca en el ejercicio de sus funciones, a través de los cauces reglamentarios.

En el Bachillerato (artículo 18 de la O.M. 25.706 de 12 de noviembre de 1992, B.O.E. del 20: corrección de erratas en B.O.E. del 27 de enero de 1993) y en la Formación Profesional Específica, el objetivo del proceso evaluador es la evaluación de los procesos de enseñanza y de la programación práctica del docente en coherencia con el logro de los objetivos educativos del currículo.

7. LA EVALUACIÓN DEL PROFESORADO UNIVERSITARIO

Como hemos indicado anteriormente, en el ámbito universitario la evaluación del profesor comienza a considerarse con la aprobación de la Ley de Reforma Universitaria y con la consiguiente elaboración de Estatutos. En la mayoría de ellos se recoge la obligatoriedad de llevar a cabo este proceso (teniendo en cuenta las opiniones de los estudiantes), siendo el Rectorado el máximo responsable. El artículo 45.3 de la L.R.U. señala:

"Los Estatutos de la Universidad dispondrán los procedimientos para la evaluación periódica del rendimiento docente y científico del profesorado, que será tenido en cuenta en los concursos a que aluden los artículos treinta y cinco y treinta y nueve, a efectos de su continuidad y promoción".

Posteriormente, con la aprobación del R.D. 1427/1986, en concreto en su artículo 8.4, se vinculó este punto al hecho de que los informes llegaran a tiempo.

Bajo este panorama legislativo, en la década de los 80, un grupo de universidades comienza a diseñar planes de evaluación cuyo objetivo era, básicamente, la evaluación de la actuación docente del profesor, tomando como método de recogida de información los cuestionarios de opinión de los alumnos (Tejedor, 1991). Con la aparición del Decreto de Retribuciones (agosto de 1989) se identifica la evaluación de la docencia y la investigación como elemento de mejora del sistema universitario, haciendo recaer la responsabilidad de la evaluación de la docencia en las universidades, mientras que el Gobierno Central se reserva la evaluación de la investigación. Posteriormente, las líneas evaluativas se plantean desde enfoques institucionales y se concretan en la puesta en marcha del Programa Experimental de Evaluación del Sistema Universitario por parte del Consejo de Universidades a partir de noviembre de 1992 (Jornet y otros, 1996). Antes de la aparición del decreto de Retribuciones, la función primordial de la evaluación del profesor era formativa. Pero, al incluir un elemento no previsto de carácter sumativo (como era el complemento salarial), los planes de evaluación diseñados no se ajustaban a este objetivo, precisando una adaptación a esta nueva situación, y una ampliación de las fuentes de información empleadas hasta entonces.

Como vemos, hemos de desarrollar planes de evaluación del profesorado en la universidad. Pero ¿qué aspectos evaluamos del profesor universitario? Tradicionalmente las funciones que se han asignado al profesor universitario han sido docencia e investigación, aunque no son las únicas, ya que también realiza tareas correspondientes a la gestión administrativa, extensión universitaria, etc. Hay que añadir, además, que la docencia no se limita exclusivamente a la actividad del aula, a la mera comunicación de saberes. El profesor debe programar su materia atendiendo a muchos aspectos como la selección del contenido a transmitir, las estrate-

gias o metodologías didácticas a emplear en el proceso de enseñanza-aprendizaje, intentando conseguir, además, el entusiasmo del estudiante por reflexionar críticamente sobre los temas desarrollados, y diseñar un justo y adecuado sistema de evaluación del rendimiento. También investigar es una tarea básica de la labor del profesor, ya que le permite profundizar en el conocimiento específico de su campo concreto de estudio, lo que repercute, además, en sus alumnos que, con facilidad, se familiarizarán con la rigurosidad y veracidad del método científico. Para conseguir todos estos objetivos, el profesor debe establecer una buena comunicación basada en la relación empática con sus discípulos, transmitiéndoles su entusiasmo por la disciplina que trate. Todas estas funciones las realiza dentro de una comunidad a la que pertenece. Por ello, la evaluación de cualquier aspecto de la labor docente requiere ser contextualizada (Escudero, 1991). Respecto a las funciones del profesor universitario, De la Orden (1987) señala:

"En primer lugar, el profesor universitario, en cuanto profesor, es una persona profesionalmente dedicada a la enseñanza, es un profesional de la educación que necesariamente comparte con los profesores de otros niveles unas funciones básicas orientadas a que otras personas aprendan. En segundo lugar, es un especialista al más alto nivel en una ciencia, lo cual comporta la capacidad y hábitos investigadores que le permitan acercarse a, y ampliar, las fronteras de su rama del saber. Estas características le asemejan al científico no vinculado a la enseñanza. En tercer lugar, es miembro de una comunidad académica, lo que supone la aceptación, y conformación de la conducta, a un conjunto específico de pautas, valores y actitudes que, de alguna manera, reflejan una determinada percepción de la realidad y caracterizan y dan sentido a una forma de vida".

Por tanto, la evaluación del profesorado debe contemplar otros aspectos tales como la evaluación de su actividad investigadora, teniendo en cuenta indicadores relacionados con la calidad, la cantidad y la incorporación de nuevas aportaciones al ámbito académico. Éste es quizá el aspecto más fomentado por la evaluación del profesor, llegando a centrarse más la labor del profesor en este ámbito en detrimento de la docencia.

En la enseñanza superior, no hemos de olvidar la evaluación de la actividad del profesor en el departamento y en otros servicios propios de la comunidad universitaria. El desarrollo de las distintas funciones del docente se realiza bajo unas condiciones académicas tales como su dedicación a la docencia, disponibilidad de medios o recursos, características del alumnado, etc.

La información que se recoge de la evaluación del profesor debe contemplar los aspectos anteriormente señalados para el desarrollo personal y profesional del docente y, consecuentemente, la mejora de la enseñanza que ofrece la institución universitaria.

Como recogen Tejedor y García-Valcárcel (1996), la Comisión Académica del Consejo de Universidades, conforme a lo previsto en el artículo segundo, apartado 3c del Real decreto 1086/89, estableció, junto a los criterios generales para la evaluación del profesorado universitario, las actividades objeto de evaluación:

a) La docencia, examinando especialmente los siguientes aspectos:
   - El nivel de cumplimiento de las obligaciones docentes, como puntualidad, asistencia, etc.
   - Claridad expositiva.
   - Actualización de los conocimientos de las asignaturas a su cargo.
   - Coordinación entre clases teóricas y actividades prácticas.
   - Congruencia del método evaluativo con los objetivos propuestos.
   - Actitud hacia los estudiantes.
   - Actualización de fórmulas pedagógicas.
   - Incentivación en la motivación del alumnado, etc.

b) Actividades de extensión universitaria, tales como:
   - Cursos de perfeccionamiento.
   - Títulos propios de universidades.
   - Otras actividades relacionadas con la docencia de programas de postgrado.
   - Otras acciones docentes institucionales propias de la universidad.

c) Actividades de servicio a la comunidad universitaria y participación en la misma.

Los diferentes sistemas de evaluación del profesorado universitario deben, por tanto, contemplar la evaluación de todos estos aspectos, evitando considerar la tarea del profesor de forma aislada, valorándola dentro de un sistema y, en su conjunto, objeto de evaluación y mejora. Si revisamos los estudios realizados sobre modelos de evaluación del profesor, comprobaremos, en demasiadas ocasiones, la ausencia de un modelo global. En la mayoría se da una especial atención a la fuente de opinión de los alumnos, vía cuestionario principalmente (constante objeto de estudio en cuanto a su validez y a los resultados que arroja), hacia la dimensión docente; y una escasa vinculación con la oferta de acciones de carácter formativo que permitan paliar las deficiencias detectadas.

En cuanto al aprovechamiento de los resultados de la evaluación del profesor, nuestra experiencia en este ámbito nos ha permitido comprobar la necesidad de su transmisión personal, a través de entrevistas a los propios interesados. En este encuentro personal, y de forma paralela a la transmisión de la cultura evaluativa de carácter eminentemente formativo por parte de la institución, el profesor aprovecha para plantear sus dudas, entender con mayor claridad los objetivos que se persiguen con el plan de evaluación, conocer la interrelación entre los distintos agentes que intervienen y el plan de formación que se diseña teniendo en cuenta los resultados globales. Si se trata de una evaluación formativa, hemos de atender a las caracterís-

ticas personales de cada docente, su trayectoria y desarrollo profesional y, en función de estos aspectos, diseñar su plan de mejora personal, no sólo con la participación en las actividades formativas diseñadas a nivel de centro, sino con la atención particular a sus planteamientos, a sus "dificultades" en el aula, a las herramientas y estrategias didácticas que desarrolla para la consecución de objetivos concretos en los procesos de enseñanza-aprendizaje... En una palabra, conocerle y ofrecerle un apoyo personalizado que le ayude al perfeccionamiento de su quehacer docente. Lo afirmado aquí para este nivel universitario, es también válido para niveles no universitarios.

En este nivel de enseñanza se ha diseñado e implantado el "Plan Nacional de Evaluación de la Calidad de las Universidades" a cargo del Consejo de Universidades. Este programa de evaluación debe contribuir, tal y como se recoge en la Guía de Evaluación (1999) a garantizar los siguientes aspectos:

a) La eficacia y eficiencia de las inversiones en Enseñanza Superior.
b) El cumplimiento de estándares internacionales de calidad en los programas de formación que permitan la movilidad y competitividad internacional.
c) Satisfacer las demandas de formación de la sociedad.
d) Responder a las necesidad de graduados que la sociedad demanda.

Se trata de un proceso de evaluación articulado en tres grandes fases. Una primera de autoevaluación, seguida de una evaluación externa, que finaliza con la elaboración de informes públicos (Informe de Evaluación de la titulación o unidad evaluada, informe de la Calidad de la Universidad, Informe sobre los resultados de la convocatoria del Plan Nacional de Evaluación de la Calidad de las Universidades).

Respecto al desarrollo de la enseñanza, este plan contempla los siguientes aspectos: atención tutorial, metodología docente (tomando como fuente de información el departamento), trabajo de los alumnos y evaluación de los aprendizajes.

En la evaluación del profesorado dentro del Plan Nacional de Evaluación se contemplan los siguientes ámbitos: la tipología del profesorado implicado en la docencia, su formación y la participación del profesorado implicado en la titulación evaluada en órganos de gobierno.

8. TÉCNICAS E INSTRUMENTOS PARA LA EVALUACIÓN DEL PROFESOR

La recogida de datos es un paso fundamental en el diseño de evaluación del profesor. Si la evaluación se basa en datos previamente recogidos, es obvio que esos datos habrán sido obtenidos por algún procedimiento que, en general, podemos denominar de medida. Sin medida, pues, no es posible hablar de evaluación, o al menos, de evaluación rigurosa y sistemática. Parece claro que, si evaluar es, en

último término valorar asignando una categoría a aquello que se evalúa, medir es, justamente, el procedimiento para definir, obtener y ofrecer información útil que posibilite la valoración posterior en la que consiste la evaluación. Los procedimientos para determinar la información apropiada a recoger, son, por tanto, un paso clave en la evaluación del profesor. Ahora bien, si las técnicas de medida no son apropiadas al objetivo específico o si el objetivo específico no refleja claramente el objetivo general, la evaluación carecerá de valor, aún cuando "técnicamente" los instrumentos estén bien construidos. También hay que prestar una atención especial a la interpretación de las puntuaciones de los propios instrumentos.

Por tanto, las técnicas se escogerán y diseñarán en función de los objetivos de la evaluación. Estas técnicas deben asegurar una recogida de información rigurosa, sistemática y controlada en orden a que los resultados de la evaluación sean fiables y válidos, y ayuden a la mejora de la labor docente y, por tanto, al aumento de la calidad de la institución.

Para la mayoría de los autores la distinción entre *técnica* e *instrumento* es importante. La técnica es el método para obtener informaciones, mientras que el instrumento es más bien el recurso concreto utilizado para obtenerlas. En este sentido, es más general el primer concepto, siendo el segundo una concreción o herramienta pedagógica de la primera. De hecho, cuando se va a programar la recogida de información de base para la evaluación, habrá que seleccionar las técnicas a utilizar y los instrumentos que se usarán para cada técnica.

Los instrumentos de recogida de información para la evaluación se pueden clasificar en cuatro tipos:

1) Aquellos en los que el agente evaluador observa y registra información directa del aula: *observación directa*.
2) Aquellos en que el agente evaluador analiza documentos y otros datos del proceso evaluador: *observación indirecta*.
3) *Encuesta*.
4) Aquellos en que es el evaluado el que reacciona a los estímulos orales o escritos presentados por el evaluador, entrevistas, cuestionarios, escalas, encuestas, etc (sean valorados por estudiantes, por colegas, por el propio profesor o por otras fuentes): *medición en sentido estricto*.

La tabla nº 1 sintetiza las técnicas e instrumentos para la evaluación del profesor, que hemos presentado con anterioridad.

8.1. La observación directa

Observar es mirar intentando prestar atención selectiva hacia determinados aspectos importantes y relevantes de una situación o de la conducta de un sujeto. Sin duda, la observación es una de las técnicas que pueden permitir al docente conseguir información directa acerca de su labor en el aula. La principal ventaja

TÉCNICAS E INSTRUMENTOS PARA LA EVALUACIÓN DEL PROFESOR		
Técnica para obtener información	***Objetivo que persigue***	***Instrumentos/Procedimientos***
1. Observación directa	El agente evaluador observa y registra información directa del aula.	– Diario profesional del profesor. – Listas de control. – Escalas de estimación. – Grabación en video. – Portfolio.
2. Observación indirecta	El agente evaluador analiza documentos y otros datos del proceso evaluador.	– Proyecto Educativo del centro. – Programación General de Aula. – Actas de reuniones. – Programas. – Trabajos de los alumnos.
3. Encuestas	Obtener información mediante la interrogación de los sujetos objeto de evaluación.	– Entrevistas. – Cuestionarios.
4. Medición en sentido estricto	Aquellos en que es el evaluado el que reacciona a los estímulos orales o escritos presentados por el evaluador, entrevistas, cuestionarios, escalas, encuestas, etc.	– Tests. Pueden ser: – Internos. – Externos.

Tabla I. Técnicas e instrumentos para la evaluación del profesor.

respecto a otras técnicas es no modificar la situación natural y permitir una interpretación global y comprensiva de una conducta o una situación.

Sabemos que la evaluación del profesorado, independientemente del método o técnica que se emplee, es compleja y difícil. Por ello, para que la observación pueda apreciar los conocimientos, la experiencia y el estilo personal del profesor, debe ser una labor sistemática. Para ello, debe contar con las siguientes características (Casanova, 1992): planificación, sistematización, objetivación, definición clara y precisa de sus objetivos, registro de los datos obtenidos mediante los instrumentos necesarios y triangulación de observadores, para contrastar y validar la información resultante.

Por tanto, para que la observación directa constituya una herramienta útil y valiosa de recogida de información, ha de ser realizada con unos objetivos claros y definidos, y debe ser convenientemente planificada, programada y desarrollada sistemáticamente.

La primera cuestión que debe plantearse el observador, como afirma Sanmartín (1978), es la determinación del campo de actuación y dentro de éste seleccionar con precisión:

— Qué es lo que se va a observar.
— Qué aspectos se van a observar detalladamente.
— Qué o cuáles sólo de un modo general o global.

Será, pues, necesario confeccionar una lista de aspectos particulares del fenómeno, conductas o tareas que se están desarrollando y que nos interesa observar. Las herramientas habituales de observación directa son: el diario profesional del profesor, listas de control, escalas de estimación, portfolio, grabación en vídeo.

a) *El diario profesional del profesor:* el objetivo fundamental del diario profesional consiste en recoger estimaciones, acontecimientos, actividades, relevantes sobre la práctica docente en el aula y las decisiones tomadas para la mejora de los aspectos que lo precisen. El repaso de anotaciones anteriores, los cambios y logros obtenidos consiguen afianzar al docente en su labor y en capacidad de modificar y perfeccionar su actuación en la clase. Se trata, por tanto, de información recogida por el propio profesor.

b) *Listas de control:* está definida por la enumeración de una serie de conductas del docente cuya presencia o ausencia se desea constatar. El observador sólo debe marcar con una cruz ("X") si el hecho se da o no durante el período de observación.

c) *Escalas de estimación:* son listas de control cualificadas, esto es, una enumeración de conductas del profesor en el aula a observar seguida de una graduación o continuum que describe la manera, forma o grado en que se manifiesta cada característica a observar. Esta técnica se viene utilizando también en alguna de sus variantes (numérica, gráfica o verbal).

La gradación puede estar formada por tres o más niveles o valores que suelen abarcar la totalidad del continuo o amplitud de la variable o conducta que se estima. Los adjetivos que expresan los distintos grados pueden apreciar descriptivamente: simplemente grado, persistencia, frecuencia, intensidad, etc.

d) *Grabación en vídeo:* permite analizar al propio docente su conducta en el aula, las relaciones que entabla con el alumnado, la metodología didáctica desarrollada en momentos concretos, etc. Proporciona un registro permanente para su visionado y examen. Posteriormente el profesor ve la grabación realizando un análisis cualitativo de su actuación en clase, detectando conductas susceptibles de mejora.

Para tratar de evitar la subjetividad de su valoración se deben tener en cuenta otras opiniones de sus compañeros. Como señala Nieto (1996): "El colega debe ser (...) alguien que a la vez sea desapasionado y crítico, pero que pueda aportar la captación de otros detalles, de otras perspectivas, de otros juicios, de modo que matice, corrija y profundice el análisis realizado por el profesor con sus únicas

fuerzas". Los principios de la técnica de la *microenseñanza* y de la *supervisión clínica entre compañeros* (Cooper, 1982) son las bases que potencian el cambio en la actuación del profesor. Posteriores grabaciones permitirán comprobar si se ha producido el cambio, o en qué aspectos ha tenido lugar y en cuáles no.

e) *El portfolio:* permite recoger documentación que contenga formas nuevas y alternativas para enseñar y evaluar los temas estudiados. El profesor recopila en una carpeta datos sobre su actuación docente, documentación que recoge las responsabilidades que tiene y puede incluir ejemplos de materiales de instrucción que ha preparado él mismo, resultados académicos de sus estudiantes, cintas de vídeo que recojan alguna de sus clases, etc.

8.2. La observación indirecta

Como hemos indicado anteriomente, consiste en el análisis de documentos que aporten información tanto sobre la labor del profesorado como de la institución educativa, lo que facilita su contextualización. Así, podemos recoger datos del proyecto Educativo del Centro, de la Programación General de Aula, actas de reuniones de los distintos miembros que conforman el centro, programas, trabajos y exámenes de los alumnos, etc. Para ello, necesitamos contar con elementos técnicos que nos permitan su análisis y valoración.

El análisis de la documentación puede ser estudiada mediante la técnica del *análisis de contenido* (Berelson, 1952). Se trata de un proceso que, siguiendo a Pérez Juste y Martínez Aragón (1989) *"presenta tres momentos fundamentales; la decisión sobre la que se tomará como unidad de análisis, la identificación, selección y estructuración de las categorías para el análisis -punto central-, y la codificación de los diversos elementos; en este punto es preciso cuidar la fiabilidad, lo que se consigue si el código de categorización está perfectamente delimitado".*

8.3. Encuestas

Los instrumentos de encuesta son, en muchas ocasiones, el procedimiento más adecuado para obtener muchas informaciones necesarias y relevantes para el proceso evaluador. Entre los principales instrumentos de encuesta podemos señalar el *cuestionario* y *la entrevista*.

a) *El cuestionario:* el cuestionario es una lista o repertorio de preguntas o afirmaciones, formuladas por escrito, a las que el alumno, profesor, etc. puede responder. Las preguntas pueden ser cerradas (elegir entre varias alternativas) o abiertas (dar una respuesta directa sin alternativas de elección).

Los cuestionarios, para que sean instrumentos útiles de recogida de información deben definir con precisión la información que es relevante, y formular cuestiones o preguntas dirigidas a recoger sólo esa información. Posteriormente, se se-

lecciona el tipo de pregunta más adecuada para cada tipo de información a recoger. Por último, se evalúa cada cuestión en particular y el cuestionario en general, al objeto de garantizar, antes de su aplicación, que recogerá, en la forma más adecuada, la información que precisamos.

b) *La entrevista:* la entrevista es un diálogo entre dos personas o entre una persona con un grupo; diálogo de carácter formal, intencionado y con objetivos claramente definidos. Es, pues, un modo directo y personal de recoger información. Requiere una gran preparación, habilidad y experiencia por parte del entrevistador si se desea conseguir, con fiabilidad y validez, datos relevantes de cara al proceso evaluador.

8.4. Medición en sentido estricto

Los instrumentos de recogida de información que hemos denominado de medición en sentido estricto reciben comúnmente el nombre tests. Se trata de instrumentos técnicamente construidos que permiten a un sujeto, en una situación definida (ante determinados reactivos o ítems), evidenciar la posesión de determinados conocimientos, habilidades, destrezas, nivel de logro, actitudes, características de personalidad, etc. Son instrumentos que permiten apreciar una variable, tal como es definida por la misma prueba o instrumento.

Los tests, a su vez, pueden ser divididos en *internos* y *externos*, con referencia explícita al lugar en que han sido construidos, dentro o fuera del centro escolar, respectivamente. Los *tests internos* son construidos por el profesor o grupo de profesores de un centro y están dirigidos específicamente a medir el logro de unos objetivos concretos, previamente definidos.

Los *tests externos* se caracterizan por (García Ramos, 1989):

a) Medir rasgos o variables de carácter general.
b) Estar construidos por empresas privadas o especialistas.
c) Estar formados por ítems o tareas previamente experimentadas en grupos normativos. Garantías de fiabilidad y validez.
d) Aplicarse y corregirse estandarizadamente, esto es, de acuerdo a rigurosas instrucciones.
e) Su aplicabilidad restringida a grupos similares o de las mismas características al grupo normativo.
f) Estar generalmente comercializados.

ACTIVIDADES PARA EL ALUMNO

1. Define indicadores de evaluación del profesorado en las siguientes áreas:
   — Planificación y organización de la enseñanza.
   — Dominio de contenidos y claridad expositiva.

   — Relaciones con los estudiantes.
   — Empleo de recursos.
   — Método de evaluación.
   — Relaciones con el equipo docente.
   — Relaciones con los otros miembros de la institución educativa.
   — Relaciones con los padres.
2. Indica los principales motivos por los que debe ser considerado el alumno, sobre todo a nivel universitario, fuente de información para la evaluación del profesorado.
3. Reflexiona sobre el uso que puede hacerse de los resultados de la evaluación del profesorado para lograr una mejor orientación e intervención renovadora en los aspectos objeto de la evaluación.
4. Describe los pasos que seguiríamos para la elaboración de una escala de observación o de un cuestionario de evaluación docente a rellenar por el alumnado.

BIBLIOGRAFÍA COMENTADA

ANGULO, F.; FERNÁNDEZ SÁNCHEZ, J. y MARTÍNEZ ARIAS, M.R. (1987). *La evaluación de la enseñanza universitaria*. Instituto de Ciencias de la Educación de la Universidad Complutense de Madrid. Madrid.

La evaluación de la enseñanza universitaria continúa presentando problemas y cuestiones pendientes a pesar de la abundante bibliografía que la sustenta. Tras reflexionar sobre las metas y objetivos de la evaluación de la enseñanza universitaria y la conceptualización de la evaluación educativa, los autores abordan el estudio de las dimensiones fundamentales en la evaluación de la calidad de la enseñanza universitaria, describiendo el proceso seguido en la elaboración de un cuestionario de evaluación del profesorado universitario, su aplicación y la valoración de los resultados obtenidos, con el objetivo de estudiar el grado de acuerdo entre profesores y estudiantes en torno a las características relevantes y/o irrelevantes para la calidad de la enseñanza.

MATEO, J.; ESCUDERO, T.; DE MIGUEL, M.; MORA, J.G. y RODRÍGUEZ ESPINAR, S. (1996). *Evaluación del Profesorado de Secundaria*. Barcelona: Cedecs.

Los autores de esta obra abordan el controvertido tema de la evaluación del profesorado, desde una doble perspectiva. Por una parte, exponen los enfoques y modelos más representativos de evaluación empleados y, por otra, analizan la opinión de los propios profesores sobre su evaluación, pretendiendo detectar las actitudes que manifiesta el Profesorado de Educación Secundaria sobre la implantación de un

sistema evaluativo de su labor profesional y conocer las características que creen debe tener el modelo y los procedimientos a utilizar en dicha evaluación.

VILLA, A. y MORALES, P. (1993). *La evaluación del profesor. Una visión de los principales problemas y enfoques en diversos contextos*. Bilbao: Servicio Central de Publicaciones del Gobierno Vasco.

La evaluación docente para la mejora del profesorado en particular y del centro en su conjunto, es el enfoque que recoge la obra, aunque también subraya los diferentes propósitos para los que puede emplearse la evaluación del profesor. Tras examinar los distintos modelos de evaluación del profesorado que se llevan a cabo en diversos contextos y realizar un análisis del hecho educativo en nuestro ámbito, se exponen una serie de conclusiones de gran utilidad para los centros y profesores que deseen evaluar sus propios proyectos y diseño educativos. También aporta, en el ámbito de la evaluación docente, un gran abanico de procedimientos y métodos, múltiples experiencias y contextos, y datos empíricos que permiten esclarecer los aspectos más polémicos en la evaluación del profesor.

YÁÑIZ ÁLVAREZ DE EULATE, C. (1998). *Un sistema de autoevaluación y mejora para el Profesorado de Educación Primaria*. Bilbao: Ediciones Mensajero.

Concepción Yáñiz realiza en esta obra un somero análisis de las funciones y tareas asignadas al profesorado de primaria en la actualidad y ante la innovación educativa, resaltando la necesidad de fomentar el trabajo en equipo del profesorado. Destaca fundamentalmente la idea de que la evaluación docente, la formación profesional permanente y la profesionalización del profesorado constituyen los pilares básicos del desarrollo profesional. También presenta un cuestionario de autoevaluación para el profesorado de primaria: PAD-Primaria (Plan de Autoevaluación Docente), instrumento útil para la recogida de información sobre la práctica docente, que facilita la evaluación de la misma. Finaliza con una reflexión sobre las diferentes fases que han de seguirse en la implantación de un plan de intervención para la mejora de la tarea educativa.

FOROS NACIONALES E INTERNACIONALES DE DISCUSIÓN Y DIRECCIONES ÚTILES EN RELACIÓN A LA EVALUACIÓN DEL PROFESORADO

FOROS/ORGANISMOS	INFORMACIÓN PROPORCIONADA
AIDIPE (Asociación Interuniversitaria de Investigación en Pedagogía)	• Asociación Interurniversitaria de Investigación en Pedagogía. Es miembro de la European Educational Research Asociation. Organiza, cada dos años, un Seminario Nacional con un tema centrado en torno al cual giran las ponencias, comunicaciones y mesas redondas. Esta asociación publica: • Revista de Investigación Educativa (RIE), con carácter semestral. • Actas de sus Seminarios. También elabora la revista: • RELIEVE: Revista Electrónica de Investigación y Evaluación Educativa [ww2.uca.es/RELIEVE].
Asociación Española de Pedagogía	Esta asociación publica la revista BORDÓN.
HEURESIS	Revista Electrónica de Investigación en la que se recogen trabajos que puedan aportar ideas y perspectivas que enriquezcan el debate sobre los problemas educativos. Aborda numerosos temas entre los que se contempla el de la calidad de la enseñanza y de los centros educativos y el de la evaluación institucional. [www.uca.es/HEURESIS].
Instituciones Internacionales de evaluación educativa	• ACER: The Australian Council for Educational Research. • U.S. Department of Education. • NCES: National Center for Education Statistics. • DPD: Direction de la Programmation et du Développement. • OCTO: Center for Applied Research on Education (University of Twente). • IIE: Instituto de Innovaçao Educational.

CAPÍTULO V

EVALUACIÓN Y GESTIÓN DE LA CALIDAD DE LOS CENTROS EDUCATIVOS

Jesus Miguel Muñoz Cantero[1]
Universidad A Coruña

OBJETIVOS

1. Asumir las diferencias conceptuales de evaluación subyacentes de las diversas corrientes evaluativas
2. Aproximarnos al concepto de modelo como paso previo a la realización de una evaluación y su implicación en el desarrollo de una teoría
3. Comprender y reflexionar sobre las distintas clasificaciones de los modelos de evaluación y analizar su incidencia en la investigación evaluativa, así como su importancia en los procesos evaluadores en los centros educativos
4. Conocer, analizar y reflexionar sobre las últimas tendencias en evaluación de centros
5. Conocer la aportación de diversos autores a la investigación evaluativa
6. Reflexionar sobre la metodología de la investigación en evaluación de centros, de su proceso
7. Reflexionar sobre la utilidad de los instrumentos para la recogida de información

CONTENIDOS

Evaluación de los centros educativos

1. Concepto y características de la evaluación
2. Modelos de evaluación de centros educativos
    - 2.1. Modelos de evaluación de organizaciones
    - 2.2. Modelos de evaluación desde un punto de vista epistemológico y metodológico
    - 2.3. Modelos de escuelas eficaces vs. de Gestión de Calidad Total.
3. Plan de evaluación de centros.
    - 3.1. Objetivos y criterios de evaluación
    - 3.2. Modelos de evaluación.
    - 3.3. Procedimientos de evaluación.
    - 3.4. Técnicas para la recogida de datos.

Actividades para el alumno
Bibliografía comentada

[1] Facultad de Ciencias de la Educación. Campus Elviña, 15071. Área de Métodos de Investigación y Diagnóstico en Educación. Universidad de A Coruña. Tfno: 981167154. Fax: 981167153. E-mail: munoz@udc.es

EVALUACIÓN DE CENTROS EDUCATIVOS

Evaluación y calidad educativa son dos conceptos que, en los últimos años, están tomando un cariz especialmente relevante en la investigación evaluativa. La necesidad de lograr una óptima calidad en la enseñanza, tanto universitaria como no universitaria, se ha convertido en un objetivo esencial y prioritario para la Administración educativa, para el conjunto de la sociedad y, como no para la propia institución educativa. En este contexto, hablar de evaluación y, por tanto, de calidad, supone hablar de fines, objetivos y criterios, de teorías y modelos, de todos los componentes que concursan y participan en el sistema educativo (Muñoz et al., 1995). Resulta evidente y lógico, que la evaluación, entendida desde esta perspectiva sugiere una apertura a un concepto global que, como señala De Miguel (1994), ha de aplicarse a todo el proceso educativo: alumnos, profesores y al centro en su totalidad.

Por esta razón se están desarrollando en nuestro país innumerables esfuerzos en atender a los procesos de evaluación como medio de reflexión y perfeccionamiento de la actividad humana que se realice, sea a nivel personal o profesional (Tejedor, 1999).

1. CONCEPTO Y CARACTERÍSTICAS DE LA EVALUACIÓN

El concepto de evaluación ha ido variando a lo largo de nuestra historia moderna. En un principio la concepción era más mercantilista, es decir, se evaluaba un producto con el fin de mejorar una serie de condicionantes para convertirlo en más eficaz. Esta idea, comercial, tuvo su repercusión en el sistema educativo europeo y, mientras en Inglaterra se aplicó al ámbito educativo con cierto éxito al tener en cuenta los procesos, en nuestro país, se tradujo rápidamente evaluación igual calificación, o dicho de otras palabras, solamente se tenía en cuenta la evaluación del aprendizaje de los alumnos. Esta concepción, ha ido cambiando en las últimas décadas, y la preocupación por considerar que la evaluación debe tener en cuenta también los procesos ha hecho que las investigaciones en este campo sean muchísimas. Por este hecho, son también numerosas las definiciones que sobre evaluación existen.

Algunas se centrarán en determinar hasta qué punto se alcanzan los *objetivos educativos* (los estudios de Tyler -1950- son un claro exponente de ello); otras hacen se centrar en la concepción de que evaluar es proporcionar información para la *toma de decisiones* (en esta línea tenemos los estudios de Cronbach -1963-, Stufflebeam -1969- y Alkin, 1969); también la evaluación ha sido considerada como la valoración del *mérito o valía* (Scriven -1967- y Eisner, 1979), o ya en la década de los 80, tenemos que la evaluación es entendida como *una actividad* que comprende, por un lado, la *descripción* entendida como recogida de datos y, por otro, el *juicio crítico*, es decir, son los valores, normas sociales y preferencias de los indivi-

duos quienes financian la evaluación (Guba y Linconln, 1982). En los noventa la evaluación es entendida como mejora (De Miguel, 1994; Mateo, 1995; Pérez Juste, 1995; Álvarez, 1988; López Mojarro, 1999).

Nevo en 1986 y, posteriormente, en una obra traducida al español en 1997[2], señala una serie de cuestiones que se deben tener en cuenta a la hora de conceptualizar la evaluación y que vamos a resumir a continuación:

— **Definición**: para este autor la evaluación hay que entenderla como la *recogida sistemática* de la información referente a la naturaleza y a la calidad de los *objetos* educativos.
— **Objetos**: los objetos educativos a evaluar serán todas aquellas entidades definibles, en este sentido podremos evaluar los programas educativos, el currículo, los centros docentes en general, el personal y el alumnado.
— **Tipo de información**: no sólo interesa obtener información de los resultados de la evaluación, o los objetivos educativos, o aquellos aspectos más referidos a la toma de decisiones, como veíamos en las definiciones anteriores, sino que es necesario examinar y obtener información de las necesidades, de los objetivos, de las estrategias, de los procesos de implementación y del medio sociopolítico.
— Pérez Juste (1995), nos señala que cada modelo de evaluación exigirá determinadas informaciones.
— **Criterios**: Nevo hace referencia a cinco criterios para juzgar el mérito o la valía: la consistencia interna, la comparación con objetos alternativos, los estándares acordados previamente, normas y valores sociables aceptables y juicios críticos de expertos.
— Si hacemos hincapié en este tema, De Miguel (1994) entiende que los criterios de evaluación deben estar unidos a los principios que orientan la actividad del centro, los objetivos concretos que se pretenden alcanzar con esos principios y los criterios de evaluación[3]. Para este autor, el criterio de evaluación *es un objetivo de calidad respecto al cual puede ser valorado un fenómeno* o un aspecto concreto del centro educativo. Reflejamos en la Tabla 1, cómo se pueden coordinar estos principios, objetivos y criterios extraídos de la obra del autor mencionado.

En esta línea López Mojarro (1999) señala que un centro docente tiene unas intenciones claramente definidas por la legislación vigente, por sus propios proyectos y por los idearios que le imprimen carácter. Además tiene unos objetivos establecidos, siendo éstos la justificación última de la existencia del centro.

[2] La obra original es "School based Evaluation: A dialogue for school improvement".
[3] Este modelo está adaptado, como De Miguel pone de manifiesto, de la obra de John (1991). A framework for examining state recourse management strategies in higher education. *Journal of Higher Education* 62 (3), 263-287.

PRINCIPIOS	OBJETIVOS	CRITERIOS PARA LA EVALUACIÓN DE RESULTADOS
Igualdad.	Acceso para sujetos en desventaja.	Porcentaje de alumnos según niveles económicos y étnicos de los centros.
Calidad.	Mejorar la calidad de los estudiantes.	Resultados de los estudiantes, evaluación comparada entre los centros.
Desarrollo económico-social.	Promover el crecimiento del capital humano.	Niveles de educación alcanzados.

Tabla 1. Modelo de coordinación de principios. Objetivos y criterios para la evaluación de resultados.

Para Pérez Juste (1997) los criterios a los que Nevo se refiere son simples referencias a la hora de valorar. Considera el criterio como la *norma para juzgar* el valor de algo o, dicho de otro modo, aquel aspecto que desde ese algo es valorado como correcto o adecuado.

En suma, los conceptos siguientes: principios del centro, objetivos y criterios se perfilan como elementos necesarios a la hora de definir la evaluación.

— **Funciones**: son numerosas las funciones que la evaluación puede cumplir. Las concepciones más clásicas van desde la evaluación centrada en la mejora (formativa) a aquella centrada en la selección, certificación o rendición de cuentas (sumativa). Otras concepciones se tienen como objetivo el aumentar la motivación (esta función es más de carácter sociopolítico pues la evaluación es utilizada para motivar y ganar apoyo público) o ejercer la autoridad (una función administrativa). En todo caso como señala Nevo, la función básica de la evaluación será la de ayudar a comprender la naturaleza del objeto de evaluación y de su calidad o, como señala López Mojarro (1999), la función de la evaluación es la mejora.

— **Audiencias**: podemos considerar a las audiencias como todos los posibles colectivos relacionados con el objeto a evaluar, así como todos aquellos que por una u otra razón tenga un interés en los resultados. Es necesario que el evaluador satisfaga las necesidades de los interesados.

— **El proceso de evaluación**: se podrían sugerir distintos procedimientos en función del planteamiento del que partamos. Siguiendo la definición de Nevo dada anteriormente podemos señalar los siguientes estadios o fases: a) comprender el problema objeto de evaluación, b) planificar la evaluación, c) recoger los datos, d) analizar datos, e) informar sobre los resultados y f) proporcionar recomendaciones; entendiendo este proceso como cíclico en donde cada fase está sujeta a ser repetida a lo largo del estudio evaluativo.

— **Métodos de Investigación**: de todos es conocida la problemática existente entre la metodología cuantitativa y la cualitativa. Lo que nos debe preocupar no es si debemos utilizar una u otra, sino qué metodología da respuesta al problema por mí planteado y cuál responde mejor a mis expectativas de éxito. Esto nos obligará, a aplicar el principio de complementariedad metodológica a fin de poder contrastar los datos con técnicas variadas y complementarias.

— **Tipos de evaluadores**: independientemente del tipo de evaluación de la que partamos, sea interna (autoevaluación, heteroevaluación o coevaluación)[4] o externa; inicial, procesual o final; global o parcial; sumativa o formativa, es necesario que los evaluadores estén formados en las competencias siguientes: a) comprender el contexto social, organizativo y personal de la evaluación así como las características específicas del objeto de evaluación; b) poseer destrezas comunicativas que le permitan establecer relaciones de empatía con los grupos humanos con los que trabaja; c) poseer conocimientos en métodos y técnicas de evaluación.

— **Estándares de evaluación**: el Comité Conjunto de Estándares de Evaluación en Educación desarrolló un conjunto de estándares divididos en cuatro grupos: estándares de utilidad (para que la evaluación satisfaga las necesidades prácticas de información); estándares de viabilidad (asegurarán que la evaluación es real y prudente); estándares de propiedad (asegurarán que la evaluación cumple con los preceptos legales y se desarrolla éticamente) y, por último, estándares de precisión (asegurarán que la evaluación es técnicamente adecuada).

Esta serie de consideraciones nos obligan a considerar el concepto de evaluación desde un punto de vista global, es decir, que se aplica a todo el proceso educativo y que tiene en cuenta algo más que los procesos de enseñanza aprendizaje, es decir, al centro en su totalidad. La evaluación no es una suma de partes sino que es, como dice Scriven (1991), una nueva disciplina con un cuerpo propio de conocimientos. No es sólo un instrumento imprescindible de ayuda para detectar lo que está ocurriendo en un centro educativo, sino que nos pone en situación de qué es lo que *necesita para mejorar*. Evaluar es conocer con la intención de mejorar, es conocer los hechos y los factores que los condicionen, es conocer de manera sistemática (López Mojarro, 1999). Y esta sistematicidad es la que la convierte no sólo en un instrumento sino que es, como señala Ruiz (1998: 18), un "proceso de análi-

[4] La evaluación interna es la llevada a cabo por los mismos miembros del centro evaluado y, podemos considerar tres tipos de evaluación interna:
Autoevaluación: evaluación realizada sobre el trabajo de los propios evaluadores.
Heteroevaluación: valoración por parte de un evaluador distinto de las personas evaluadas.
Coevaluación: distintos sectores del centro se evalúan mutuamente.

sis estructurado y reflexivo, que permite comprender la naturaleza del objeto de estudio y emitir juicios de valor sobre el mismo, proporcionando información para ayudar a mejorar y ajustar la acción educativa". De Miguel (1994), dentro de esta línea, va un poco más allá y considera que la evaluación no se relega al final del proceso sino que es inherente al proceso, lo cual hace que la evaluación se convierta en parte integrante del aprendizaje, la enseñanza y la gestión de los centros educativos. Los resultados de la evaluación se convierten en los elementos que ayudan a definir y redefinir los objetivos y a estimular a profesores y a alumnos a tomar responsabilidades en el aprendizaje, la enseñanza y la gestión de los centros donde estudian y trabajan.

La evaluación así entendida, como proceso, reúne una serie de *características*, que Rotger (1989) clasifica en:

- La evaluación debe estar perfectamente integrada dentro del currículo vitae, de manera que ésta sea una fase más del proceso.
- Debe tener un componente formativo. Este carácter se imprimirá tanto en el proceso educativo como en los resultados.
- Ha de ser continua. Esto supone que los efectos de la evaluación han de ir conociéndose a lo largo de todo el proceso, no al final. A fin de realizar las correcciones pertinentes para alcanzar el mayor número posible de objetivos fijados en un principio.
- Ha de ser recurrente. A través de la dimensión retroalimentadora que proporciona se mejorará el proceso en función de los resultados obtenidos (está en íntima relación con la característica anterior).
- Criterial, en relación a unos objetivos prefijados al inicio y que conduzcan el proceso en dirección a los mismos.
- Decisoria. Esto significa que la toma de decisiones ha de permitir mejorar el proceso y los resultados del aprendizaje.
- Cooperativa: en tanto que la evaluación atañe a todos los estamentos en ella implicados, éstos deben participar en su proceso.
- Cardona (1994) y Ruiz (1998) añaden a estas características su carácter indirecto (dado que las variables en el campo educativo sólo pueden ser medidas en tanto que son observadas), científico (en sus instrumentos, metodología y análisis de datos) y referencial (pues su fin es referenciar los logros obtenidos con los objetivos propuestos).
- Como consecuencia directa de esta visión de entender la evaluación ha sensibilizado de modo significativo a todos los agentes implicados en el hecho educativo, entendiendo que la educación es un bien público que es importante mejorar y cuidar; con lo que la administración y los poderes públicos han llevado a cabo una nueva ordenación del sistema educativo, incrementándose las inversiones públicas en el campo educativo. En concreto en temas de calidad, de eficacia y de eficiencia del sistema educativo,

en general, y de los centros, en particular. Lo que ha motivado que el concepto de evaluación educativa abarque a los alumnos, a los profesores, centros y, en conclusión, al sistema educativo en su conjunto (Hernández Pina, 1995).

2. MODELOS DE EVALUACIÓN DE CENTROS EDUCATIVOS

Para afrontar este punto hemos optado por dar una visión de los modelos de evaluación desde tres ópticas diferentes. Pretendemos obtener, por un lado una visión temporal de los diversos modelos desde el punto de vista de las organizaciones. Posteriormente, los analizaremos desde un punto de vista epistemológico y metodológico. Finalmente, haremos una revisión de los modelos actuales de evaluación partiendo del criterio de calidad.

2.1. Modelos de evaluación de organizaciones

Hoy día se tiene asumido que los centros educativos hay que tomarlos como organizaciones en la medida que poseen una estructura compleja y que no puede ser analizada desde enfoques parciales. Las investigaciones actuales se centran en analizar la cultura, personalidad o funcionamiento propios desde puntos de vista metodológicos diferentes.

Toda teoría sobre las organizaciones nos remite a un proceso de evaluación que permite, por un lado, a los gobiernos, conocer si las instituciones educativas/organizaciones escolares, cumplen o no con sus cometidos, a la vez de si los recursos que les han sido asignados son empleados *eficaz* y *eficientemente*. Por otro lado, la evaluación permite comprobar a las propias instituciones si se han alcanzado los objetivos propuestos y detectar las disfunciones internas que impiden obtener resultados más satisfactorios. Es decir, se comprueba la eficacia y se identifican aquellos elementos de disfuncionalidad y susceptibles de cambio (evaluación para el cambio).

Cuando hablamos de **modelos** de evaluación referida a las organizaciones, lo que se está postulando es, en palabras de De Miguel (1989: 30) y Mateo (1995: 75), una **teoría** sobre su estructura y funcionamiento que especificamos a través de variables relacionadas entre sí y en conexión con otras que estableceremos como indicadores[5] de eficacia y calidad. Determinar los *criterios de calidad o eficacia* que deben asumir las organizaciones constituye *el punto central de toda evaluación*,

[5] Entenderemos por indicador tal como lo hace De la Orden (1995) como una variable, significativa, frecuentemente cuantitativa, que tiene carácter normativo. De Miguel (1994) nos advierte que los indicadores señalan la presencia o ausencia de un hecho y que frecuentemente tienen una expresión numérica que recibe el nombre de índice.

por lo que deberemos remitirnos al *enfoque teórico* desde el que se conceptualiza. En palabras de estos autores, la relación entre teoría y evaluación es tan estrecha que podemos entender la existencia de tantos modelos de evaluación como enfoques teóricos se formulen.

Esto nos puede llevar a pensar que toda evaluación sobre las organizaciones posee un modelo. No debemos olvidar que el modelo es una guía, es una representación ideal y abstracta de una realidad basada en referentes teóricos. En tanto representación teórico-hipotético de una realidad debe poseer ciertas características: a) debe estar basado en una teoría; b) las variables con las que trabaje deben ser susceptibles de ser especificadas operacionalmente; c) el modelo debe poseer validez empírica y d) se deben validar y generalizar las aplicaciones del modelo.

Para realizar una clasificación de éstos podíamos tomar diversos criterios. Vamos a optar en primer lugar por realizar reseñar la clasificación presentada por De Miguel (1989)[6] puesto que su visión nos va a permitir analizar cuál fue el pasado y cuál es el futuro de la investigación de las instituciones educativas. Posteriormente, daremos algunas pinceladas de los principales modelos de evaluación existentes partiendo de otro criterio.

2.1.1. *Modelos centrados en los resultados*

Su fin era el de evaluar principalmente la eficacia de las instituciones y organizaciones pero desde la perspectiva de los resultados (perspectiva externa) expresados en términos cuantitativos, utilizando diseños de input-output y técnicas correlacionales de análisis. Dentro de esta línea se encuadran la mayoría de los trabajos evaluativos centrados en la eficacia. Estimar el rendimiento de una organización conlleva a evaluar los resultados en función de las metas propuestas (eficacia).

Entre las ventajas que presentan este tipo de modelos es que permiten proporcionar datos sobre causas y efectos, llegando a identificar las características de los centros pero no a discriminar cuáles eras las variables que influían en la eficacia. Entre las desventajas podríamos señalar que no permiten evaluar la valiosidad de las metas, imponen controles difíciles para el contexto educativo y tienen serias limitaciones conceptuales y metodológicas (Fernández, 1990).

[6] Aunque el esquema que vamos a seguir es el que presenta De Miguel en un trabajo publicado en la "Revista de Investigación Educativa" en 1989, este autor realiza otra clasificación más reduccionista que la expuesta en su obra de 1994: "Evaluación para la calidad de los institutos de secundaria" y publicada por Escuela Española, en la que resume los modelos en dos grandes bloques: modelos centrados en la eficacia (en el que siguiendo a Scott -1981- Shortell y Kaluzny -1983- analiza las instituciones a partir de la estructura, los procesos y las conductas desde tres enfoques: racional, natural y abierto) y modelos centrados en la evaluación del cambio.

Entre los trabajos que podemos citar dentro de esta línea tenemos los de Averch y col. (1972); Brookover y al (1974), Rutter (1979, 1983); Madaus y al. (1979); Purkey y Smith (1983); Lotto y Astuto, 1984…

Los criterios o indicadores de éxito de eficacia normalmente utilizados son:

- Evaluación en función de los objetivos: las escuelas eficaces son las que cumplen con los objetivos previstos.
- En relación a los recursos: el criterio de evaluación de las escuelas eficaces radica en la capacidad para adquirir, generar y utilizar recursos con el fin de producir más y mejor.
- Indicadores de carácter social: los criterios de los que se parte es la contribución y/o servicios que las organizaciones dan a la sociedad. Los indicadores dependerán de las directrices políticas sobre la función de la escuela en cada momento.

2.1.2. *Modelos centrados en la eficiencia de los procesos internos*

Los modelos centrados en este punto de vista ven la evaluación organizacional como un proceso de eficiencia desde una doble perspectiva. Por un lado eficiencia en el funcionamiento interno -estructura y operaciones de la organización-. Por el otro, la satisfacción de los miembros de la organización. Evaluar el rendimiento de una organización desde la eficiencia supone estimar dichos resultados en función de sus costos.

Son modelos centrados en la eficiencia interna de la escuela (notemos que hay un giro respecto a los modelos anteriores cuyo objetivo era la eficacia), lo que supone un interés por aspectos internos de la propia organización. Hay un interés por analizar factores psicosociales como la estructura organizativa, el director, el clima y los procesos del aula. Su objetivo es analizar el funcionamiento y la satisfacción de los miembros del centro. Responden a una metodología explicativa con diseños input-context-output (utilizando análisis de regresión) y diseños comparativos.

Entre las ventajas que presenta es que toman en consideración todas las variables. Por el contrario siguen teniendo ciertas limitaciones conceptuales y metodológicas.

Entre los autores que podríamos señalar dentro de esta línea tenemos a Goodman y Pennings (1977); Van de Ven y Ferry (1980); Hall (1987); Shoemoker y Fraser (1981).

Estos modelos toman tres indicadores como referencia:

- Estudios sobre el papel del director y el staff: toda evaluación de organizaciones no puede prescindir de los aspectos prácticos relativos a la dirección y administración de la organización. Estos modelos ven en el director la persona responsable de la eficacia de la organización en tanto que influye en la consecución o no de las metas, clima escolar, eficiencia instruccional…

- Sobre la eficiencia instruccional: estos modelos utilizan como indicadores aquellas variables implícitas en la dinámica del aula como las oportunidades que tiene el alumno para aprender, la distribución del tiempo, la interacción profesor-alumno, currículum propuesto…
- Centrados en el clima institucional: estos modelos parten del hecho de que los procesos internos que se dan en una organización tienen como consecuencia una relación entre sus miembros que tiene una implicación con su grado de satisfacción personal y, por tanto, de eficacia de la organización. Son indicadores de estos modelos: los procesos de integración y cohesión internos, las percepciones, actitudes y expectativas de los docentes y estudiantes, formas de comunicación…

2.1.3. *Modelos causales*

Si analizamos los modelos anteriores, unos intentan explicar las organizaciones desde una perspectiva externa (análisis de resultados) y otros desde una perspectiva interna (cómo funciona la organización). Sin embargo, resulta sumamente difícil prescindir a una de las variables procesos y, a la otra, de las diferencias.

Con el fin de superar estas limitaciones, los modelos causales, nacen como una respuesta integradora que nos permite analizar las relaciones de causa y efecto entre las variables relativas al proceso y las relativas a los resultados. Lo que pretenden estudiar estos modelos es determinar el patrón de causación de la eficacia del centro, a través de análisis causales, análisis de estructuras de covarianza y análisis factoriales confirmatorio-exploratorio.

Este tipo de estudios tienen la ventaja de permitir comprobar las relaciones causa-efecto, si bien si existe una clara dificultad de evaluar los procesos.

Algunos autores que podemos citar en esta línea tenemos a Yorke (1977); Stöel (1987); De Miguel (1986); Fuentes (1988); Smylie (1988).

2.1.4. *Modelos culturales*

El hecho de considerar a las escuelas como sistemas sociales obliga a considerar los aspectos culturales que están implícitos en una organización. Estos modelos parten de perspectivas ecológicas, fenomenológicas, etnográficas,… centrando su análisis en aspectos como la comunicación, los pensamientos, las interacciones, imágenes sobre el centro, dentro de lo que se conoce como paradigma emergente.

Los modelos culturales los podemos clasificar en tres grandes bloques:

- La cultura es considerada como elementos de información acerca de los individuos y los grupos dentro de una población.
- La cultura es considerada como estructuras conceptuales y simbólicas a partir de las cuales construyen la realidad de la escuela los integrantes de la misma.

- La cultura es considerada como un conjunto de propósitos y significados diferenciales generados dentro de un grupo concreto en su relación con el entorno social y político.

2.1.5. *Modelos centrados en la evaluación del cambio*

A la par que se realizaban las investigaciones sobre las organizaciones eficaces, se desarrollaban otros trabajos cuyo fin era la innovación y cambio.

Estos modelos buscan la mejora de la escuela partiendo del hecho de que todos los implicados en su actividad pueden introducir cambios en los procesos y procedimientos con el fin de mejorar su organización y funcionamiento interno.

Estos modelos son considerados como la alternativa hacia la cual han de tender los procesos evaluativos de los centros educativos teniendo como objetivo averiguar aquellos aspectos de los centros que deben mejorar y cómo pueden hacerlo. Tienen, pues, una función de control (se controla la eficacia o consecución de los objetivos), diagnóstica (de aquellos elementos que producen disfuncionalidad) y de intervención (con el fin de lograr la mejora). Desde la teoría organizacional el éxito viene dado por:

- Por un lado, por el número y calidad de los productos u objetivos alcanzados y,
- Por otro, por los procesos de mejora de la efectividad y calidad de vida dentro de la organización.

Estos modelos han supuesto cambios a nivel conceptual como metodológico y entre los que de Miguel (1994) señala los siguientes:

- Desde un punto de vista conceptual se pasa de los modelos causales a otros de carácter contextual. Esto supone que de diseños que buscan la relación causa-efecto se pase a enfoques orientados al cambio y cuyo fin es la de analizar las relaciones de los fenómenos de manera interactiva y procesual.
- Si el objetivo anterior era analizar el rendimiento de los alumnos, ahora el fin es el de diagnosticar los problemas internos de la institución para proceder a resolverlos. El proceso se va a centrar en una serie de indicadores que permitan observar la institución.
- Los modelos para el cambio consideran a cada institución como única, con su propia autonomía. Lo que le confiere el derecho de poner en marcha sus propios procesos de mejora. Es decir, parte de la *autorreflexión* o *autoevaluación*.

La teoría actual sobre la evaluación de las instituciones tiende hacia una evaluación externa-social para comprobar su eficacia y, hacia una evaluación interna para identificar aquellos aspectos disfuncionales de los centros. Se trata de diseños orientados hacia la mejora de los centros escolares partiendo de enfoques procesua-

les en donde la autoevaluación es el eje vertebrador y en los que el evaluador no tiene por qué ser un elemento externo a la institución. Éste es el enfoque propuesto por la OCDE (1991) y que actualiza el CERI (1994) de cara a la consecución de escuelas eficaces.

2.2. Modelos de evaluación desde un punto de vista epistemológico y metodológico

Como hemos podido apreciar, las distintas formas de concebir el éxito de los centros han dado lugar a la aparición de diversos modelos o paradigmas de evaluación. Es muy difícil encontrar en la realidad modelos puros que puedan ser encajados en una determinada concepción puesto que todos ellos tienen su forma de entender qué es la evaluación, cuál es su ámbito y objeto y cómo puede llevarse a cabo. Desde una perspectiva metodológica se ha tendido a señalar la existencia de dos grandes grupos de modelos[7]: por una parte tenemos los modelos clásicos orientados hacia la consecución de metas (tal es el caso de las propuestas tylerianas) o los orientados hacia la toma de decisiones (modelo CIPP de Stufflbeam). Por otra, los modelos alternativos, tal es el caso del modelo alternativo de Stake, Illuminativo de Hamilton, crítica artística de Eisner y el modelo de evaluación democrática de McDonald.

Vamos a profundizar un poco más en estos modelos partiendo de la clasificación de Colás y Rebollo (1993) y Medina y Villar (1995), tomando como base el desarrollo de Stufflebeam y Shinkfield (1987). Esta clasificación distingue tres momentos: modelos eficientistas u objetivistas, modelos de corte humanístico, fenomenológicos o subjetivistas y, modelos holísticos[8].

2.2.1. *Modelos eficientistas-conductistas*

En este tipo de modelos encajarían aquellos que De Miguel (1989) señalaba como los centrados en los resultados, en los modelos centrados en la eficiencia y

[7] Popham (1980) clasifica los modelos de evaluación en: modelos de consecución de metas, modelos de enjuiciamiento que acentúan criterios intrínsecos, modelos de enjuiciamiento que acentúan criterios extrínsecos y modelos de facilitación de decisiones.

De la Orden (1993, 1995b) distingue dos modalidades o concepciones de modelos de centros escolares según hagan referencia o no al producto. Los que tienen en cuenta el producto se estructuran de acuerdo a una serie de variables seleccionadas en función de su relación con el rendimiento de los alumnos. Los que no hacen referencia al producto no los considera como modelos diagramables. Así entre los primeros distingue: modelos input-output, modelos procesuales de eficacia institucional, modelos procesuales con énfasis en la clase, modelos sistémicos.

[8] Pilar Colás y Ángeles Rebollo en su obra de 1993: "Evaluación de programas" realizan una clasificación muy similar: modelos subjetivistas, modelos objetivistas y modelo crítico.

los modelos causales. Su preocupación se centra en la valoración de los objetivos y rentabilización de los resultados.

Lo que pretenden este tipo de modelos es controlar las variables (con diseños evaluativos de investigación experimental, de tipo científico-tecnológico) que inciden en los resultados para lograr el mejor producto, aprovechando al máximo los recursos.

Características:

- Evaluador: generalmente externo. Actúa como técnico especializado.
- Técnicas o instrumentos: cuestionarios cerrados, test estandarizados, pruebas objetivas,… (métodos cuantitativos). Se busca la objetividad.
- Debilidades: no contempla las interacciones humanas. Los centros educativos no permiten un total y riguroso control experimental.
- Autores: Suchman, E.A. (1967); Tyler (1967); Cronbach, L.J. (1980); Stufflebeam, D.L. y Skinfield, A.J. (1987); Pophan, W.J. (1981)

Modelo de evaluación basado en los objetivos de Tyler. Tyler consideraba que la evaluación debe determinar la congruencia entre trabajo y objetivos. Para él el proceso de evaluación es aquel que determina hasta qué punto los objetivos educativos han sido alcanzados mediante los programas de currículos y enseñanza. Para ello procuró establecer, clasificar y definir los objetivos en términos de rendimiento.

Este modelo está orientado hacia la toma de decisiones, las cuales deberán estar en consonancia con los objetivos y los resultados obtenidos, utilizando pruebas de valoración para medir el logro de los objetivos previstos, el ritmo personal de evolución de los individuos, y la adecuación a parámetros o normas fijadas en el ámbito sociopolítico de que se trate (Medina y Villar, 1995: 47).

Aunque este modelo ha sido criticado sobre todo a que la elección de objetivos quedaba restringida a los objetivos de rendimiento, con lo que la evaluación tendía a convertirse en algo incompleto, es de destacar la enorme importancia que este autor da a la evaluación para juzgar el proceso de un programa como sus resultados finales, poniendo las bases del primer método sistemático de evaluación (Stufflebeam y Shinkfield, 1987).

El procedimiento evaluativo de las ocho etapas de Metfessel y Michael. Desarrollado en 1967, estos autores pensaban al igual que Tyler que la valoración de los resultados podría hacer más perfectas las decisiones educativas, desarrollando un paradigma evaluativo de ocho etapas destinado a ayudar al personal de los centros educativos a evaluar la consecución de objetivos. Éstas son:

- Implicación de todos los miembros de la comunidad escolar en la evaluación.
- Hacer una lista priorizada de metas y objetivos.
- Traducir los objetivos específicos de rendimiento de manera que puedan ponerse en práctica para facilitar el aprendizaje.

- Seleccionar los instrumentos que permitan comprobar la efectividad de los procesos educativos a fin de analizar si se cumplen los objetivos.
- Aplicar los instrumentos para analizar si se producen cambios en el rendimiento respecto a los objetivos seleccionados.
- Analizar los datos a través de métodos estadísticos.
- Interpretación de los datos.
- Hacer recomendaciones a las audiencias correspondientes.

Estos autores hacen responsables de la evaluación a evaluadores profesionales, estudiantes y profesores en cada una de las etapas y en distinto grado. También insistían en que aquellas decisiones que supusiesen algún enjuiciamiento deberán incluirse en todas las etapas de proceso evaluativo a fin de que los implicados en cada fase realicen sus reajustes a tenor de la información recibida.

Modelo de planificación educativa de Cronbach. Este modelo surge por el interés del autor en evaluación de programas de ciencias sociales, considerando que las evaluaciones están concebidas para cumplir una función política, por lo que hay que estimular la *planificación* de las mismas.

El evaluador no debe ser único, por lo que la responsabilidad de la evaluación la debe tener un equipo que, por un lado, reparta prioridades y responsabilidades entre ellos y, por otro, detalle su planificación interna. Aunque la evaluación suele ser realizada a petición de los administradores, para que un evaluador externo aplique técnicas de muestreo, valoración y análisis, los evaluadores deberán esforzarse por llegar a los ciudadanos de tal manera que clarifique los problemas de los implicados y proporcione información (clara, oportuna, exacta, válida y amplia) y sugerencias para la resolución de problemas. Para lograrlo el evaluador deberá poseer una mentalidad abierta y dotes de buen comunicador y deberá atender a las Unidades (cualquier individuo o clase), a los tratamientos (que inciden sobre las unidades) y a las operaciones de observación (técnicas de recogidas de datos, análisis de datos, etc…).

Para Cronbach la evaluación debe ocuparse en algo más que en los objetivos, debe considerarse como un estudio tanto de proceso como de los resultados. Pone también énfasis en cómo los resultados obtenidos en un campo determinado pueden utilizarse para la toma de decisiones a otros campos.

Modelo C.I.P.P: Stufflebeam. Si el modelo de Tyler se desarrolla a principios de los 40 y el de Cronbach a finales, el modelo CIPP surge a finales de los 60. Este modelo supone una nueva definición de evaluación distinta a la de los autores precedentes: "la evaluación es el proceso de identificar, obtener y proporcionar infomación útil y descriptiva acerca del valor y mérito de las metas, la planificación, la realización y el impacto de un objeto determinado, con el fin de servir de guía para la toma de decisiones, solucionar los problemas de responsabilidad y promover la comprensión de los fenómenos estudiados" (Stufflebeam y Skinkfield, 1987: 183).

Esta definición presenta a la evaluación no como una prueba sino como un proceso. Este modelo parte de que lo esencialmente importante de una evaluación no es demostrar sino perfeccionar. La utilización del modelo CIPP (Contexto, Input, Proceso, Producto) está concebido para promover el desarrollo y ayudar a los directivos y personal responsable de la institución a obtener y utilizar una información continua y sistemática.

Los contenidos a evaluar se estructuran en cuatro grandes dimensiones:

— Evaluación de Contexto (C): su objetivo es determinar el contexto institucional, la población objeto a estudio, diagnosticar necesidades y problemas, etc.

— Evaluación de Input (I): su objetivo es identificar la capacidad del sistema, estrategias, planificación de procedimientos, presupuestos y programas,…

— Evaluación de Proceso: identifica o pronostica durante el proceso los defectos de la planificación, de su realización,…

— Evaluación de Producto (P): recopilar descripciones y juicios acerca de los resultados y relacionarlos con los objetivos y la información de las otras tres evaluaciones anteriores, interpretando su valor y mérito.

— Los instrumentos a utilizar son muy variados: análisis de sistemas, inspección, audiciones, entrevistas, test, inventarios, análisis de recursos humanos y materiales, visitas, etc…[9]

Método evaluativo centrado en la figura de Stake[10]. Este método que fue denominado "modelo de la figura para la evaluación educacional", entendiendo por figura la imagen de la evaluación educativa, se desarrolló a finales de los años 60 y está construido sobre lo que Tyler pensaba de los evaluadores en tanto que éstos deben comparar los resultados deseados y los observados, destacando la importancia de analizar la consecución de los resultados en función de los objetivos planteados.

Urge a los evaluadores y educadores para que presten más atención a todo el conjunto de la evaluación que incluye la descripción (descripción de la evaluación, de su complejidad y de su importancia), juicio (una evaluación no tiene sentido hasta que se emite un juicio), esquema para la recopilación de datos (antecedentes -información relevante-, transacciones de la enseñanza -encuentros de los estudiantes con todas las personas-) y los resultados (lo que se consigue), análisis de las congruencias (¿se han cumplido los propósitos?) y contingencias, identificación de

[9] Una de las más notables aplicaciones en nuestro país la tenemos en el Plan de Evaluación de Centros (Plan EVA) que el Ministerio de Educación y Ciencia puso en marcha a partir de 1990.

[10] Para algunos autores como Worthen y Sanders (1987), (Colás y Rebollo, 1993) a este autor lo inscribirían dentro de una corriente más de tipo humanista, subjetivista, dado que rompió con muchos de los tópicos tylerianos al desarrollar a principios de los 70 una nueva serie de conceptos y que presentó como "Evaluación respondente".

las normas pertinentes y a menudo opuestas y, las múltiples utilizaciones de la evaluación.

2.2.2. *Modelos humanísticos, fenomenológicos o subjetivistas*

En este tipo de modelos existe una creciente preocupación por los individuos, por sus actividades, por los procesos y fenómenos que producen. Incluiríamos aquí los modelos culturales de los que nos hablaba De Miguel (1984), dado que, al preocuparse por los procesos tienen en cuenta variables como: los valores de las personas, sus creencias, el sustrato cultural de su propia vida, sus actitudes hacia el entorno, etc…

Su fin es mejorar los procesos educativos y de aprendizaje con la perspectiva de su orientación a los resultados. Utilizarán tanto los instrumentos cualitativos como cuantitativos.

El evaluador colaborará con los implicados en los procesos evaluativos, estableciendo un clima de intercomunicación, de responsabilidad y de reflexión orientado a una mejora personal y profesional.

El método contrapuesto de evaluación de Owens. Este método se desarrolla a finales de los setenta. Supone ya un paso hacia el fin de la tradición tyleriana. Su fin último es perfeccionar la toma de decisiones, para lo cual, será necesario que el método proporcione a los que toman esas decisiones información para que elaboren proyectos y juzguen los resultados previstos y los no previstos para que puedan actuar.

Fue desarrollado para dar una explicación a aquellas variables que son menos cuantificables o tangibles (clima escolar, actitudes de los alumnos,…) y que, normalmente, son rechazadas. Los procedimientos de contraposición se desarrollan como una audiencia administrativa destinada a juzgar los méritos de un caso en el que están implicadas dos partes opuestas (dos grupos de evaluadores, por ejemplo) de forma que se puedan tomar las mejores decisiones de una forma abierta y participativa, al poder intervenir aquellas personas que están interesadas en el proceso evaluativo (Stufflebeam y Shinkfield, 1987: 296).

Owens sugiere siete modos principales en los que puede llevarse a cabo el modelo de contraposición para satisfacer los propósitos de una evaluación educacional: exploraciones de los valores del currículum existente u otro nuevo (se desarrollaría a través de un debate abierto sobre la pertinencia o no de un currículum concreto), selección de nuevos libros de texto (se podría celebrar una audiencia entre profesores que están a favor y otros en contra de determinados libros de texto), estimación de la congruencia entre una innovación y el sistema existente (la toma de decisiones puede ir orientada hacia la pertinencia de esa innovación), revelación de las distintas interpretaciones hechas por distintos representantes de los mismos datos (en el centro escolar están profesores, alumnos, administrativos,…

La toma de decisiones puede ser colectiva, como si de un jurado se tratara), información a los profesores, directores y administradores, resolución de disputas acerca de los contratos de trabajo y, por último, llegar a la decisión que debe ponerse en práctica.

Para Owens estos métodos de contraposición pueden ser relevantes para la aplicación de los modelos desarrollados por Stufflebeam y Stake.

El evaluador sigue considerándose como externo pero que sirve de mediador en el debate entre los implicados en la discusión evaluadora.

El modelo judicial de Wolf. El modelo de Robert Wolf fue desarrollado en 1973 a partir del de Owens. Es una variante que ofrece la forma de juicio. Se basa en un tribunal ante el que se ofrecen argumentos y justificaciones para facilitar un acuerdo.

Para Wolf la solución a la evaluación convencional no reside en recopilar gran cantidad de datos técnicos sino en la clarificación de esos datos y en su información. Los evaluadores pueden centrarse en un determinado problema y contar con el testimonio de otras personas y de otros evaluadores; pueden analizar los métodos mediante los que se recogen y analizan los datos,... con el fin de obtener un punto de vista más equilibrado.

Las etapas de este modelo son cuatro:

- — La etapa que propone programas principalmente a través de entrevistas.
- — La etapa que selecciona los problemas. Se reducen los problemas hasta que se identifica aquellos que sirvan para una audiencia. Un grupo de investigadores decidirán la importancia de cada problema.
- — La etapa de preparación de argumentos: los equipos evaluativos preparan entonces los argumentos formales.
- — La etapa de la audiencia: con dos partes, una previa en la que se exponen los argumentos de ambos equipos y se fijan las normas de procedimiento y otra que es propiamente la audiencia.

Modelo de atención al cliente de Scriven. Para algunos autores (Farley y otros, 1985; Rebollo y Colás, 1993) este modelo es considerado dentro de los modelos objetivistas. Medina y Villar (1995) lo incluyen dentro de los modelos fenomenológicos. Stufflebeam y Shinkfield, (1987) lo incluyen dentro de los modelos holísticos, por lo que parece que no hay una unanimidad a la hora de clasificar este modelo. Nosotros lo vamos a poner a caballo entre los centrados en los objetivos y los puramente holísticos.

Este planteamiento surge como respuesta opuesta a los modelos basados en objetivos logrados de Tyler, dando más importancia a aquellos logros no previstos que los que figuraban como previstos.

Desarrollado por Scriven en 1967 se centra más en analizar cómo afectan los programas educativos a las personas que en las metas, es decir, descansa sobre el análisis de las necesidades del cliente, pasando los objetivos a un segundo plano y

emita un juicio de valor bien informados sobre un objeto basándose en las pruebas acumuladas procedentes de su comparación con otros objetos, a fin de satisfacer las necesidades del consumidor. La evaluación es comparativa.

Scriven propone que el evaluador, aunque en las primeras etapas del proceso evaluador reconoce la necesidad de la "evaluación amateur" -autoevaluación- son necesarios evaluadores profesionales externos, desconozca cuáles son los objetivos previstos para que éstos no obstaculicen la evaluación de otros aspectos, lo que haría que ésta fuese incompleta y errónea (lo que se ha denominado evaluación sin metas). Si bien, este proceso es reversible, puesto que una vez analizados todos los efectos inesperados se puede empezar el proceso basado en las metas para ver si éstas han sido alcanzadas.

Otorga a la evaluación dos funciones: la formativa, que ayuda a desarrollar el proceso evaluativo ayudando a la planificación y al perfeccionamiento de cualquier cosa; y la sumativa, que calcula el valor del objeto una vez desarrollado, proporcionando juicios acerca de hasta qué punto las metas reflejan, de una manera válida, las necesidades que han sido valoradas.

Crítica artística de Eisner. Su concepción de evaluación parte de una línea humanístico-crítica porque el hecho de que las personas puedan emitir sus opiniones hace que se tenga una visión más equilibrada, completa y creativa sobre el objeto de evaluación. Por esta razón su finalidad radica en que las personas implicadas en procesos evaluativos posean actitudes interpretativas y valorativas del proceso educativo y de su contexto. Para ello, el evaluador que actúa como promotor de opiniones de profesores, alumnos,... que le permitan conocer la realidad educativa, interpretarla y juzgarla teniendo en cuenta el contexto, los procesos educativos, los procesos emergentes, las relaciones entre los diferentes procesos,... buscando la valoración crítica que le conduzca a la emisión de juicios de valor.

2.2.3. *Modelos holísticos*

Los modelos adscritos a esta tipología empiezan a desarrollarse a partir de los años 60 coincidiendo con la aparición del paradigma interpretativo, de corte eminentemente cualitativo.

Estos modelos parten de una visión global de los diferentes componentes del centro educativo, lo que obliga a tener en cuenta las opiniones, valores, creencias y aptitudes de todos los implicados en él. Desde esta perspectiva la evaluación es concebida un proceso de comprensión y valoración de los procesos y de los resultados. El evaluador es un cooperador con los participantes de la evaluación.

Evaluación respondente de Stake. Fruto de la evolución del "modelo de la figura" visto anteriormente, Stake desarrolla, a mediados de los 70, una alternativa evaluativa que rompe con la tradición tyleriana. Este método subraya marcos para el desarrollo del aprendizaje, las transacciones didácticas, los datos para los juicios, el informe holístico y la asistencia a los educadores.

El fin de este tipo de evaluación es ayudar a comprender a profesores, alumnos, personal de administración,... los problemas, virtudes y defectos de los programas educativos. Para ello el evaluador debe hacer referencia a las opiniones de la gente sin juzgarlas, recopilando información a lo largo de todo el proceso y dando respuesta a esos problemas o cuestiones que se le planteen. Para lograrlo, las planificaciones serán abiertas y flexibles tendiendo más a la descripción que a la valoración.

La metodología incluirá primero la observación para posteriormente intervenir (investigación naturalista) introduciendo técnicas como el estudio de casos, los muestreos intencionados, informes narrativos, etc utilizando, si es necesario, otros métodos.

La comunicación con el cliente ocurrirá de manera natural, frecuente e informalmente, puesto que el fin es localizar continuamente cuestiones que le preocupen.

Stake identifica cuatro partes en la estructura sustancial de la evaluación respondente: los problemas (áreas de desacuerdo, cuestiones a resolver...), un esquema para la recopilación de los datos (que procederá del modelo de la figura), los observadores humanos (como los mejores instrumentos de investigación) y la validación (a través de mucha información que le permita tener una visión del objeto a estudio).

El método requiere gran cantidad de observación, en un proceso en el que no existen las fases lineales y donde la información continua es lo más importante. Esta estructura funcional de la evaluación responde a la forma de evaluación reloj: las doce, hablar con los clientes y audiencias interesadas; la una: identificación del alcance del problema entre las partes interesadas; las dos: identificar las actividades; las tres: descubrir los propósitos de la evaluación y los intereses de la gente; las cuatro: analizar las cuestiones y problemas; las cinco: identificar los datos necesarios para investigar los problemas, las seis: seleccionar los observadores, jueces e instrumentos; las siete: observar los antecedentes de las transacciones y los resultados propuestos; las ocho: desarrollar temas, preparar descripciones y estudiar casos concretos; las nueve: validación; las diez: esquematización de la información para las audiencias; las once: reunir los informes formales si los hay.

La evaluación iluminativa de Parlett y Hamilton. Frente a los modelos evaluativos convencionales de tradiciones experimentales o psicométricas de alcance muy limitado surgen otros métodos, cuyo origen está en la antropología social, que quieren analizar la realidad como un todo, utilizando para ello la descripción y la interpretación (frente a la valoración y la predicción); es lo que se ha dado llamar metodología iluminativa. Los autores quieren "iluminar" los distintos componentes de la evaluación para describirlos y hacerlos más comprensivos en relación con el contexto.

La adopción de una evaluación iluminativa supone entender dos aspectos: la descripción del sistema de enseñanza y el medio de aprendizaje como el ambiente en donde trabajan alumnos y profesores.

Es un método que no intenta manipular, controlar o eliminar las variables situacionales sino que, por el contrario, utilizando una combinación de técnicas intenta analizar y comprender las relaciones entre las creencias y las prácticas y entre los modelos de organización y los de los individuos, poniendo énfasis en la observación de las aulas, entrevistas a profesores y estudiantes, cuestionarios y test, y fuentes documentales y sobre antecedentes. Las estrategias metodológicas emergen de procesos de negociación.

El evaluador promoverá el debate entre los participantes para que de ahí afloren opiniones, interpretaciones y juicios de valor, actuando como cooperador.

Evaluación holística de MacDonald. Para MacDonald la evaluación debe tener en cuenta y debe ser motivo de su preocupación todos los datos posibles y el contexto, puesto que, la acción humana dentro de las instituciones está expuesta a multitud de variables que influyen en ella.

El método holístico implica analizar cada acto educativo dentro de la totalidad, implicando a los protagonistas en la interpretación de la realidad en la que aparecen concepciones, creencias, valores e intereses que deben ser explicados a través de la participación colaborativa en el debate, convirtiéndose éstos en el contenido de la evaluación. A partir de aquí, desde esta perspectiva naturalista, la evaluación se propone facilitar y promover el cambio, desarrollando actividades de análisis y valoración (McDonald, 1983; Medina y Villar, 1995).

El evaluador ofrece información y procedimientos de búsqueda y contrastación, a través del diálogo, la discusión, la búsqueda y el análisis,... (Colás y Rebollo, 1993). Para promover la interpretación se acudirá a técnicas cualitativas.

2.3. Modelos de Escuelas Eficaces vs. de Gestión de Calidad Total

Numerosos trabajos centrados en buscar cuáles son las características de las escuelas eficaces o escuelas con éxito (effective schools), que respondían que un buen centro de calidad dependía de sus profesores y alumnos (buenos profesores + buenos alumnos = buen rendimiento).

En los años 70-80 los investigadores (Brookover et al., 1979; Madaus, Airasian y Kellaghan, 1980; Cremers y Scheerens, 1989), tratan de aislar otra serie de factores que permitan distinguir entre un centro eficaz de uno ineficaz tomando como criterio el rendimiento medido a través de pruebas estándares. En este sentido De Miguel (1994), analizando el trabajo de Edmonds y colaboradores (1978) identifica los cinco factores que presentan mayor correlación con la eficacia de una escuela:

"— Liderazgo del director y atención que presta a la instrucción.
— Grandes expectativas de los profesores sobre los alumnos.
— Énfasis del trabajo en el aula sobre las habilidades básicas.
— Control continuo del progreso del alumno.
— Clima ordenado y seguro en el centro".

Purkey y Smith (1983), presentan una serie de indicadores que hacen que las escuelas sean deseables por su eficacia o excelencia agrupándolos en dos dimensiones: variables organizativas y estructurales (liderazgo pedagógico, clima ordenado y seguro, altas expectativas, autonomía curricular, desarrollo del personal de la escuela, máxima dedicación al tiempo de aprendizaje, mejora de la instrucción, participación de los padres, estabilidad del profesorado) y variables de proceso (visión de lo que se desea, seguimiento de la comunidad e implicación en un proyecto común, objetivos, importancia de los resultados académicos, trabajo en equipo).

Fullan (1985), en su preocupación de cómo funcionan las organizaciones y cómo influyen en la calidad introduce nuevos conceptos de proceso, como comunicación e interacción, trabajo colaborativo, consenso,...

Scheerens y Stoel (1988) proponen un modelo multiconceptual de escuelas eficaces, en el que las variables que definen un centro educativo constituyen condiciones para el establecimiento a nivel de aula de medidas instructivas relacionadas con el rendimiento eficaz (Gento 1996).

Otros trabajos se centraron en el análisis del contexto como elemento que influye en numerosos factores que inciden en la eficacia (Firestone y Herriot, 1982; Stoel y Cheerens, 1988; Kyle, 1985).

Posteriormente Scheerens (1990, 1992) desarrolla un modelo ecléptico que tiene en cuenta relaciones multinivelares entre (p.e. el contexto influye en las políticas del centro), los efectos causales intermedios (el efecto del liderazgo instructivo influye en los alumnos a través de los profesores) y las relaciones recíprocas (altos niveles de rendimiento aumentan las expectativas de los profesores).

Poco a poco, de un movimiento centrado en la eficacia y de los factores que contribuyen a la misma, se pasa a una nueva visión de evaluar los centros educativos, a un movimiento de mejorar la escuela: incidir en la transformación y cambio de los centros educativos.

En esta línea se encuentra el movimiento patrocinado por la OCDE (1991) y que actualiza el CERI (1994), que intenta incidir en la transformación y cambio de los centros educativos. Este movimiento considera el centro como centro de cambio, el lugar donde hay que plantear las reformas educativas. Este movimiento hace hincapié en:

- La adecuación de las reformas educativas a las características de cada centro.
- El establecimiento de procesos de mejora en las metas u objetivos del centro atendiendo a profesores, alumnos y padres. Lo que significa que no sólo se ha de atender al rendimiento de los alumnos.

La mejora de la institución educativa pasa por:

- Reformar las condiciones internas referidas a los procesos de enseñanza-aprendizaje.

- Reformar las condiciones internas referidas a la organización y funcionamiento del centro.
- Preocupación por el clima educativo.
- Toda transformación lleva implícito un proceso de planificación, implementación y evaluación.

La OCDE considera clave en la consecución de las escuelas eficaces las siguientes áreas: el currículo (planificación, aplicación y evaluación), el papel vital de los profesores, la organización de la escuela, los resultados y la evaluación. Álvarez (1998) se centra en tres aspectos para señalar las diferencias existentes entre el movimiento de las escuelas eficaces y el modelo de calidad y que nosotros reflejamos en la Tabla II.

	Sentido estático vs. dinámico	Tipo de Liderazgo	La relación proceso-producto
Escuelas Eficaces	Más importancia es la oferta. Eficacia: alto nivel de éxito académico. Indicadores: hacen referencia a resultados. Normalmente estándares nacionales. Evaluación externa.	Liderazgo claro y fuerte de estilo personal y jerárquico. Predomina el principio de autoridad.	Objetivo: resultados académicos. Producto: éxito en la instrucción.
Modelo de Gestión de Calidad	Parte de la realidad del centro y su contexto. Los centros son dinámicos. Autoevaluación: detecta puntos fuertes y áreas de mejora. El centro responde a las necesidades de sus clientes que siempre son dinámicas y cambiantes.	Liderazgo horizontal (cada responsable es líder de su proceso). Desarrolla liderazgo compartido.	Igual importantes procesos que resultados. Los alumnos aprenden si se implican en los procesos. Identifica cinco tipos de resultados: financieros, académicos, organizativos, de servicios, de interface.

Tabla II. Diferencias entre el movimiento de escuelas eficaces y el modelo de calidad.

Es necesario, pues, antes de introducir cualquier aspecto evaluativo, realizar un análisis de los factores que inciden en el centro, considerado como organización. En este contexto, como señala De Miguel (1994), surge *la evaluación institucional como alternativa para cambiar las organizaciones educativas*. Cuyo fin primordial es facilitar procesos de autorregulación, es decir, autoevaluación.

En esta línea se mueven casi todos los programas que actualmente se han puesto en marcha en nuestro país y concretamente, el puesto por el MEC a través de

Modelo Europeo de Gestión de la Calidad para la evaluación de centros no universitarios o universitarios, basados todos ellos en una combinación de evaluación Externa (realizada por la Inspección o Comités de Expertos en caso de la enseñanza Universitaria) e Interna (iniciada por el Equipo Directivo de los centros).

En resumen, actualmente, la gestión de Calidad se apoya en la generación de un cambio cultural en las organizaciones, promovido por la orientación permanente de las organizaciones hacia la mejora de procesos y resultados.

2.3.1. *Modelos de calidad*

A continuación vamos a reseñar una serie de modelos de estimación de la calidad de una institución partiendo de la clasificación dada por Gento (1996), por considerar con él, que son los que mayor interés tienen para las instituciones educativas.

El modelo de control de calidad global de la empresa. El modelo[11] aunque nace en EE.UU influido por la concepción organizativa de Tylor (1947), se extendió rápidamente a empresas europeas. Aunque se aplicó sobre todo a empresas es fácilmente extrapolable a otros campos que en el caso educativo podría regirse por los principios siguientes:

- La calidad educativa (producto) viene determinada en la medida que satisface a alumnos, padres,... (clientes). El cliente es la prioridad absoluta.
- La gestión de los recursos humanos de los centros se fundamentará en la formación continua de los mismos (los recursos es el elemento más importante del modelo).
- En la distribución de los bienes, los proveedores (eslabón intermedio de la cadena) deben prestar su colaboración.
- La gestión y funcionamiento del centro tenderá a la prevención de errores, más que a la rectificación.
- Todos los componentes del centro educativo tienen que funcionar de acuerdo a criterios de calidad.
- La mejora ha de ser continua en cada uno de los sectores del centro.
- El centro se acomodará a las exigencias de los alumnos (clientes en general), con nuevas metodologías, técnicas, servicios,... con el fin de satisfacer sus necesidades.
- La permanente mejora irá acompañada de mecanismos de promoción interna permanente de todos (alumnos a otros grados, personal de administración,...).
- Implicación de todos los miembros de la institución en actividades y procesos de mejora: diseño, gestión, evaluación,...

[11] El modelo fue definido originalmente como "Company-Wide Quality Control" (CWQC).

El Modelo del premio Deming a la calidad. Se estableció en Japón por la Japan Union of scientists and Enginieers. Este premio es un reconocimiento a Edwart W. Deming[12]. Lo que se busca es analizar la implicación de todos los componentes de la organización. El enfoque valorativo se centra en el *control estadístico de la calidad*; en el proceso utilizado para aplicar este control.

Es un modelo que no tiene en cuenta la satisfacción del cliente. Dedica 9 de sus diez s a estimar procesos de mejoramiento de calidad y uno de ellos a la estimación de los resultados (teniendo un peso igual cada criterio: un 10%).

Estos criterios[13] con algunos s de valoración son:

- Política general: política de calidad, planificación a corto plazo y sus relaciones, comunicación interna de la política…
- Gestión de la organización: comunicación entre Departamentos, aprovechamiento personal, actividades del personal…
- Educación: planificación actividades educativas, nivel de conciencia para la calidad, la organización y control de la calidad…
- Recogida de información: exterior a la compañía, comunicación interdepartamental, procesamiento de la información (estadística)…
- Análisis: selección de problemas relevantes, métodos de análisis, técnicas especializadas.
- Estandarización: aplicación estadísticas, aplicación estándares, sistemas de estandarización…
- Control: sistemas de control de costos y calidad de productos, resultados de las actividades de control, propuestas de mejora…
- Garantía de calidad: mejoramiento procesos de control, evaluación calidad producto, inspección, mejoramiento producto…
- Resultados: proyección de resultados no medibles cuantitativamente y cuantitativos.
- Planificación futura: planificación futuras actividades, medidas para resolver áreas con problemas…

Como podemos observar también resulta fácil la aplicación de este modelo al campo educativo. Es un modelo que tendría en cuenta, igual que el anterior, toda la institución educativa en su conjunto. La diferencia estriba en los principios, criterios o s de valoración.

El Modelo del Premio "Malcolm Baldrige[14]" a la calidad. Es el equivalente norteamericano al premio Deming Japonés. Este modelo se centra en la implicación total de todos los componentes de la organización (al igual que el Deming) para valorar las empresas.

Es un modelo que tiende al *mejoramiento de la gestión* para la calidad y, vuelve a considerar la satisfacción de los clientes. El Premio a la calidad se otorga tras un proceso de tres fases: evaluación por escrito, visitas a empresas que obtuvieron alta puntuación en el examen escrito y estimación final de los resultados.

La primera de las fases, la evaluación por escrito abarca siete categorías (con porcentajes de incidencia diferentes) en cuatro elementos principales:

1. **Impulso:**
   — Liderazgo de los directivos (10%) como impulsores de la calidad.
   — Indicadores: liderazgo del Director, organización para la calidad…
2. **Sistema de mejora para la calidad, incluye:**
   — Información y análisis (7%).
   — Indicadores: enfoque y gestión de la calidad, análisis de datos e información calidad…
   — Planificación para la calidad (6%).
   — Indicadores: proceso estratégico de planificación para calidad, objetivos y planes de calidad.
   — Recursos humanos (15%).
   — Indicadores: gestión recursos, implicación personal, educación para la calidad…
   — Garantía de calidad de productos y servicios (14%).
   — Indicadores: control calidad procesos, evaluación de la calidad, documentación, calidad proveedores…
3. **Medida de progreso:**
   — Resultados de calidad (18%), comprende las mejoras que se derivan de las exigencias y expectativas de los clientes.
   — Indicadores: resultados de calidad productos, de proveedores…
4. **Meta:**
   — Satisfacción del cliente (30%), es el objetivo final del proceso.
   — Indicadores: gestión de relaciones con los clientes, satisfacción de los clientes, resultados de la satisfacción de los clientes…

Cada una de estas categorías es valorada teniendo en cuenta tres criterios:

1. Enfoque: métodos empleados para valorar los objetivos.
2. Desarrollo: amplitud con la que las áreas se aplican a las áreas más relevantes.
3. Resultados: productos y efectos en la consecución de los objetivos

Modelo de la organización Internacional para la Estandarización (ISO 9004-2). Este modelo fue creado para la valoración de las empresas productoras de servicios. Esta evaluación puede aplicarse tanto a los clientes externos como destinatarios de los servicios, como a los propios trabajadores o empleados.

[12] Edward W. Deming fue un experto norteamericano que llevó a Japón el mensaje de control de calidad.

[13] Podemos ampliar más los s de valoración de cada uno de los modelos, podemos consultar la obra de Gento Palacios, S. (1996). *Instituciones educativas para la calidad total*. La Muralla. Madrid.

[14] Malcolm Baldrige fue secretario de Comercio con el Presidente Reagan e impulsor en Estados Unidos de la Campaña Nacional por la Calidad.

Este modelo contempla la valoración de:

- Principios organizativos: responsabilidad de la gestión, estructura de la calidad, personal y recursos materiales e interacción con los clientes.
- Ámbito de operaciones a realizar: proceso de marketing, diseño de procesos, proceso de distribución de servicios, ejecución y análisis del servicio.

Modelo Europeo de Gestión de la Calidad[15]. La European Foundation for Quality Management considera la evaluación de la calidad en su totalidad. En esta línea ha desarrollado un modelo que en un principio surgió para dar los premisos de calidad a las empresas.

Este modelo ha tenido amplia repercusión en el campo educativo europeo y, como no en España, de ahí que nos centremos en él.

En el BOE nº. 131, de 2 de junio de 1998, se establece la resolución de la Dirección General de Centros Educativos por la que se dictan instrucciones para la implantación, con carácter experimental, del Modelo Europeo de Gestión de Calidad en los Centros Docentes.

La adopción del modelo de la EFQM por el Ministerio de Educación y Cultura se basa, sobre todo, en su potencial de adecuación a los entornos educativos públicos. Este modelo combina el interés por las personas con los procesos y resultados. Asociado al modelo de aplicación, va la implementación de un proceso de autoevaluación que permite valorar el progreso de la organización y establecer planes de mejora, sin perjuicio de una evaluación externa.

En el curso 96/97, se pone en marcha la primera educación del Plan Anual de Mejora y se adapta el Modelo Europeo de Gestión de Calidad a las características de los centros públicos. En el curso 98/99, se consolida la regulación sobre los planes de mejora y se generaliza la estructura del modelo de autoevaluación, como referencia para el diseño por los centros, de su evaluación interna.

A partir de aquí, el Ministerio elabora una "Guía para la autoevaluación" que viene acompañada por otras cinco publicaciones que complementan y configuran un conjunto de materiales dirigidos a los centros educativos con el fin de proporcionar las herramientas necesarias a todos los centros que quieran aplicar el modelo: "guía para la autoevaluación", "cuestionario para la autoevaluación", "formulario para la autoevaluación", "caso práctico de autoevaluación de un centro -exposición del caso-", "caso práctico de autoevaluación de un centro -valoración del caso-" y "plan anual de mejora. Materiales para el diagnóstico".

La estructura permite al centro aprender, mediante la comparación consigo mismo, y le ayuda a la planificación, en la definición de estrategias, en el seguimiento de los progresos seguidos y en la corrección de errores como efecto de su actividad.

[15] La aplicación del Modelo por el Ministerio de Educación lo podemos encontrar en la página web del Ministerio en el apartado referido a la calidad (www.mec.es).

Este modelo[16] -Figura 1- tiene en cuenta 9 criterios, 5 *agentes* (Liderazgo, Gestión del personal, Planificación y Estrategia, Recursos y Procesos) que reflejan el cómo de la gestión, y 4 *resultados* (satisfacción del personal, satisfacción del cliente, impacto en la sociedad, resultados del centro educativo) que permiten conocer y valorar lo que el centro obtiene como efecto de su actividad.

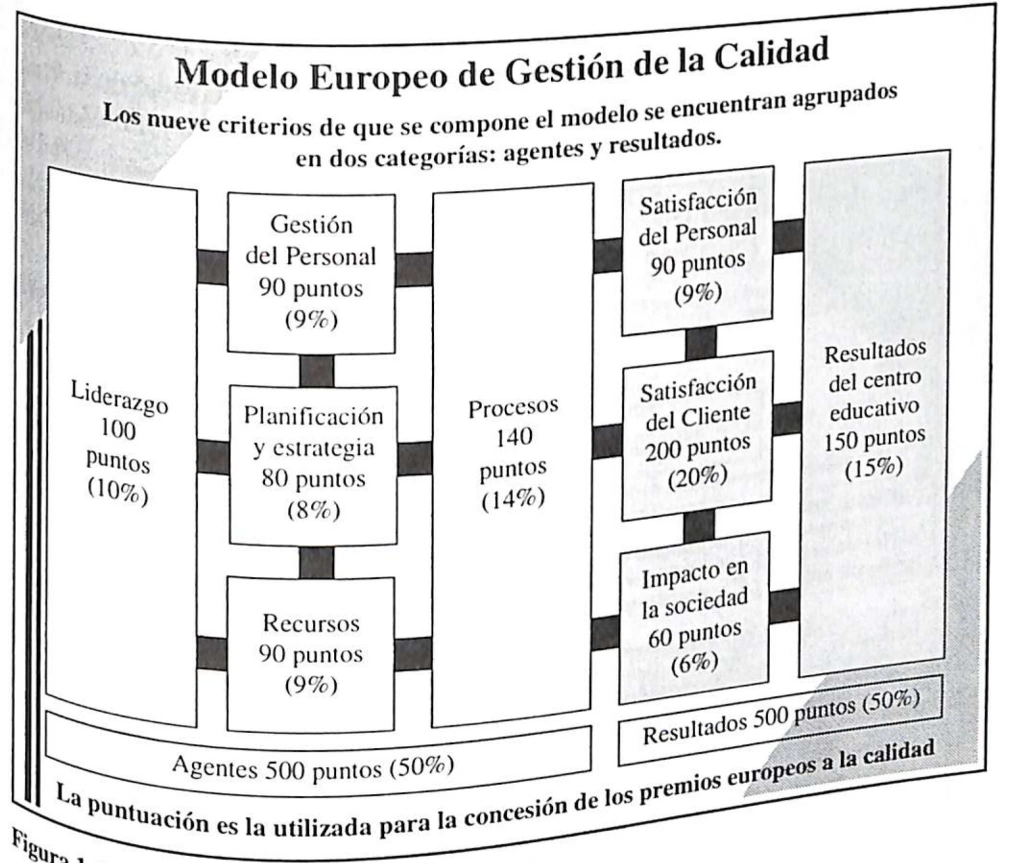


Figura 1. Modelo Europeo de Gestión de la Calidad.

[16] Los siguientes criterios vienen definidos en la "guía para la autoevaluación" del Modelo Europeo de Gestión de Calidad:
Liderazgo: se refiere al comportamiento y actuación del Equipo Directivo y del resto de los responsables para guiar el centro educativo hacia la gestión de la calidad.
Planificación y estrategia: se refiere a la Misión, Visión, Valores y Dirección Estratégica del centro educativo, así como a la forma en que éstos se implantan en los Proyectos educativos.
Gestión del personal: se refiere a cómo el centro educativo utiliza el máximo potencial de su personal para mejorar continuamente.
Recursos: se refiere a la gestión, utilización y conservación de los recursos.
Procesos: se refiere a cómo se identifican, gestionan y revisan los procesos y a cómo se corrigen.
Satisfacción del cliente: se refiere a qué consigue el centro educativo en lo relativo a la satisfacción de sus clientes.

Cada uno de estos criterios se subdivide en subcriterios, y éstos, a su vez, en áreas, con el fin de orientar al centro en su autoevaluación. Lo que se examinará será si se cumple o no lo que representa cada área y el grado de aplicación alcanzado en el centro educativo.

Así, a modo de ejemplo, en las Figuras 2 y 3 se presentan dos criterios, uno referido a la categoría de agentes y, otro, a la de resultados, con sus subcriterios y algunas de las áreas que presentan.

A través de la autoevaluación, lo que se logra es un diagnóstico de la situación del centro, de sus puntos fuertes y se detectan las áreas de mejora. El propio

Modelo Europeo de Gestión de la Calidad

CATEGORÍA: agente

- Criterio 1: Liderazgo
  - * Comportamiento y actuación del Equipo Directivo.
  - * El criterio ha de reflejar cómo todos los que tienen responsabilidad en el equipo directivo estimulan, apoyan y fomentan la gestión de calidad en tanto que proceso para la mejora continua.
- Subcategorías:

SUBCRITERIOS:
La autoevaluación debe mostrar de qué manera los líderes

Demuestran su compromiso con la cultura de la gestión de calidad	**Apoyan mejoras y la implicación de todos, ofreciendo los recursos y ayudas apropiadas**	**Se implican clientes, proveedores y otras organizaciones externas**	**Reconocen y valoran a tiempo los esfuerzos y logros de las personas interesadas en el centro educativo (clientes, personal, proveedores) u otros implicados en los resultados**
Áreas: – Fomentar la comprensión de su desarrollo. – Hacer participar a todo el personal en el proceso de calidad. – Revisar y mejorar la efectividad de su liderazgo. – Incluir compromisos y logros en la gestión de calidad. – ...	Áreas: ¿cómo actúan los líderes para?...: – Ayudar activamente a los que emprendan iniciativas de calidad. – Facilitar actividades de mejora continua. – Ayudar en la definición de prioridades en las actividades de mejora. – ...	Áreas: ¿cómo los líderes toman iniciativas positivas: – Ayudar a todos los equipos del centro a que identifiquen a sus clientes y proveedores. – Participar en relaciones de cooperación con clientes, instituciones, empresas, ...	Áreas: ¿cómo los líderes... – Reconocen las opiniones del personal. – Conocen y estimulan el trabajo de las personas y equipos. – Hacen públicos los éxitos educativos. – Premian públicamente a personas y equipos. – ...

Figura 2. Cateoría: agente. Liderazgo.

Satisfacción del personal: se refiere a qué consigue el centro educativo respecto a la satisfacción de su personal.
Impacto en la sociedad: se refiere a qué es lo que finalmente consigue el centro educativo a la hora de satisfacer las necesidades y expectativas de la sociedad en general y de su entorno en particular.
Resultados del centro educativo: se refiere a lo que consigue el centro educativo en relación a la planificación y estrategia, y con respecto a la satisfacción de sus necesidades y expectativas de los alumnos, de las familias y, en general, de los clientes.

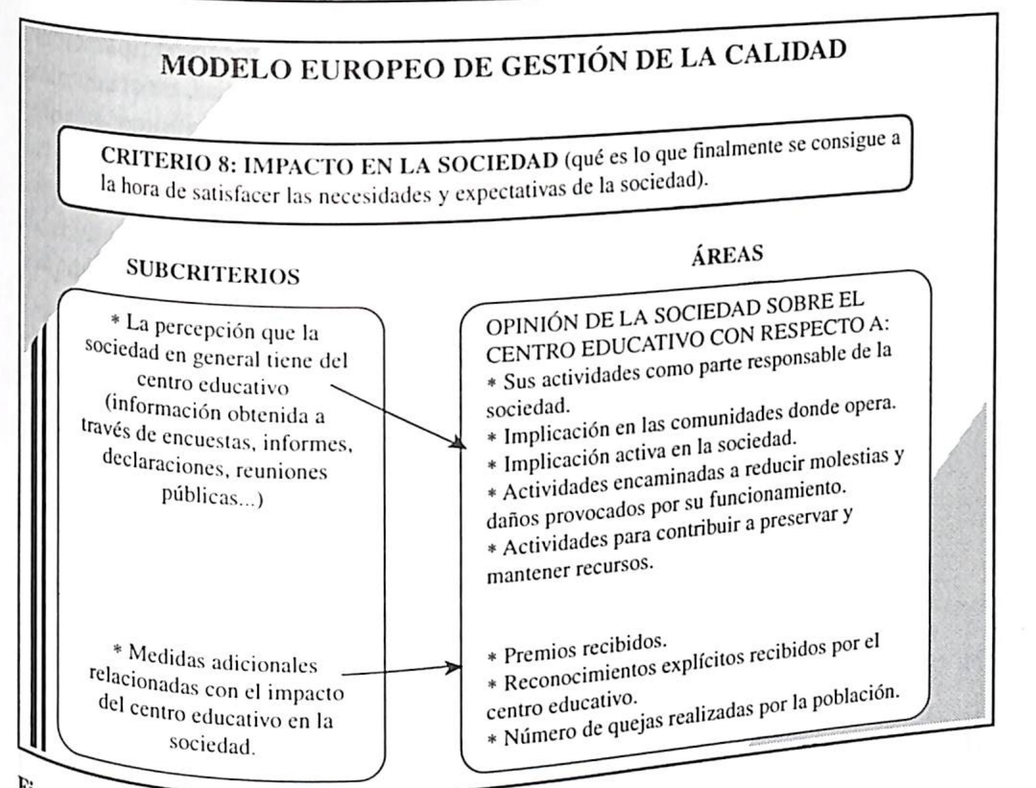


Figura 3. Categoría resultados: impacto en la sociedad.

Plan establece que, finalizada ésta, se intente contestar a preguntas como: ¿qué puntos fuertes han sido identificados y que deberán aprovecharse? ¿cuáles de éstos deben desarrollarse más aún?, y aún hay otros interrogantes, referidos a las áreas de mejora que han de ser abordadas, con las que no se trabajará inmediatamente, o a cómo se realizará el seguimiento de las acciones de mejora.

Para realizar este diagnóstico, el Modelo prevé la utilización de dos instrumentos de recogida de información:

- El formulario, que pone el acento en los conceptos de enfoque/despliegue y magnitud/alcance[17].
- El cuestionario: se compone de 67 preguntas que hacen referencia a los nueve criterios. Su fin es ayudar al centro a determinar su posición así

[17] * **Enfoque:** manera de afrontar o solucionar un problema concreto o bien alcanzar un objetivo previsto.
* **Despliegue:** extensión a la que se aplica el enfoque (extensión hacia todos los niveles y áreas de la organización, así como todos los procesos y servicios relevantes).
* **Magnitud:** es la medida de los resultados.
* **Alcance:** expresa el grado de cobertura y la relevancia de los resultados reflejados en las áreas o indicadores de un determinado subcriterio.

como las orientaciones y prioridades futuras, establecer comparaciones con los resultados de otros centros y, por último, estimular, en el centro, la cultura de la mejora continua mediante autoevaluación periódica. Realizado este diagnóstico se pueden establecer planes de mejora sobre las áreas deficitarias, además de provocar un seguimiento de estos planes. Este proceso de autoevaluación terminaría con una evaluación de los objetivos alcanzados y la detección de nuevas áreas de mejora, lo que supone un proceso de mejora continua y cíclica tal como reflejamos en la Figura 4.

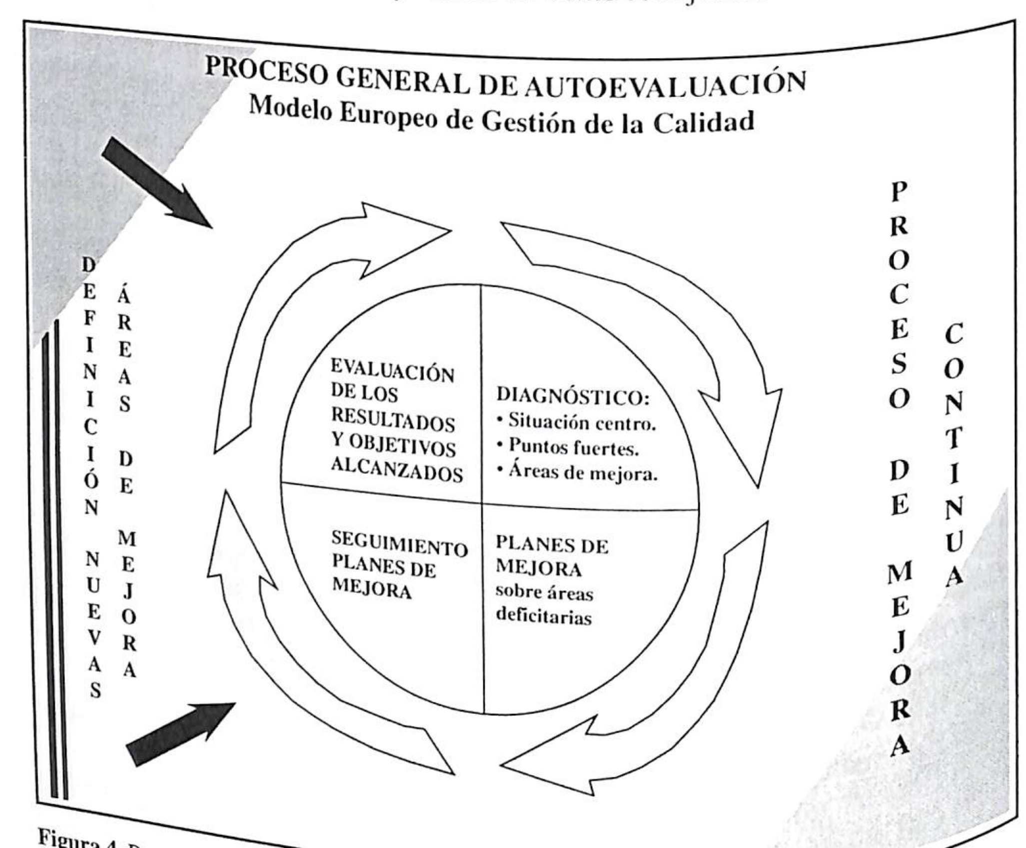


Figura 4. Proceso general de autoevaluación.

Definidos los criterios y objetivos, así como los modelos y el procedimiento de la evaluación, debemos establecer propuestas de mejora. Nos vamos a limitar en este punto, únicamente, a presentar la Orden del Plan Anual de Mejora en los Centros docentes públicos dependientes del MEC (BOE nº. 141, de 13 de junio de 1998).

Las propuestas de mejora parten de la evaluación realizada por el Equipo de Calidad del centro formado por el Equipo Directivo que ha sido formado en gestión de Calidad por la Dirección Provincial.

Este Equipo Directivo realizará un primer diagnóstico de la situación del centro a través de entrevistas con alumnos/familias/personal y, adaptando el material de diagnóstico que ha elaborado la Dirección General de Centros Educativos.

Realizado este primer diagnóstico, el Equipo Directivo (dividido en tres subgrupos) procederá a realizar la autoevaluación sobre los nueve criterios del Modelo anteriormente descritos, con el fin de determinar las áreas de mejora críticas o las asumibles por el centro.

A partir de aquí, presentarán al Claustro propuestas de planes de mejora y, posteriormente, al Consejo Escolar, que determinará los que deben ser desarrollados, remitiéndolos al equipo provincial respectivo. Los pondrán en marcha los *equipos de mejora* que se constituirán para tal fin, y que, además, ejecutarán el Plan de Mejora (Figura 5).

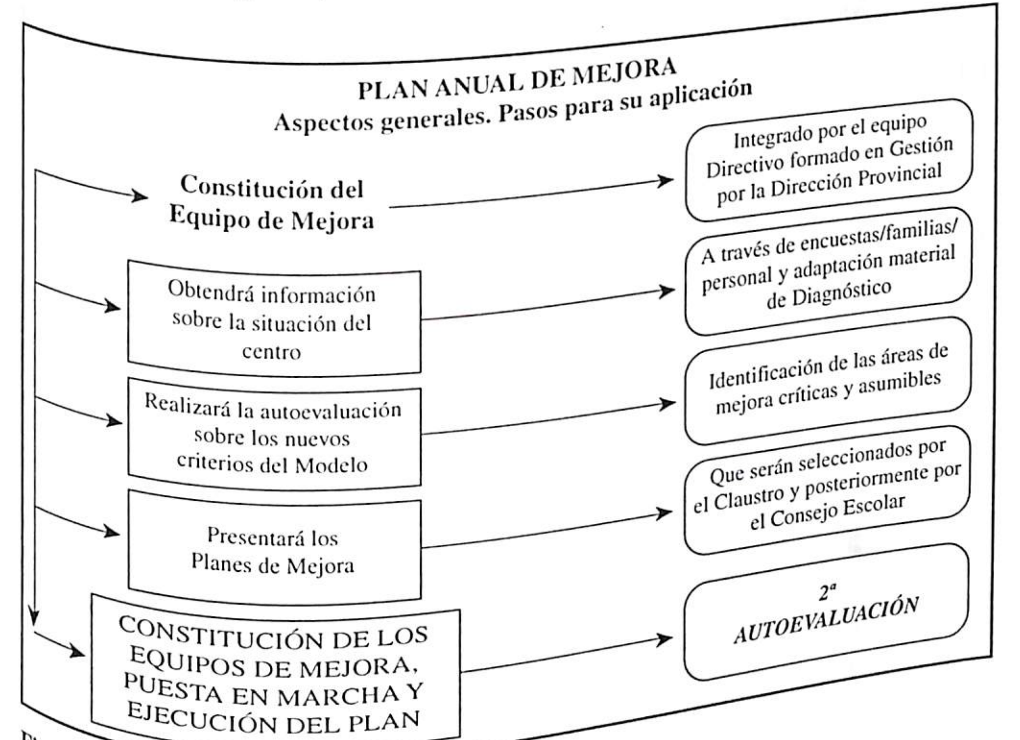


Figura 5. Aspectos generales del Plan anual de mejora.

Este plan anual de mejora se configura como un instrumento para aprender, como organización, como una herramienta para mejorar la gestión educativa, como un medio para elevar la calidad de los centros, y como un compromiso del centro y la Administración.

Sus principales características tal como son recogidas en el documento publicado por el Ministerio de Educación y Cultura (BOE nº. 141) son:

— "Debe estar precedido de un diagnóstico explícito de la situación de partida del centro con relación al área o áreas prioritarias sobre las que se centrará el Plan.
— La identificación de las áreas de mejora ha de ser objetiva y debe apoyarse en hechos o en resultados antes que en juicios subjetivos o meras apariencias.
— Los objetivos de mejora han de ser realistas, concretos, evaluables y alcanzables en un curso escolar, sin perjuicio de que puedan integrarse como parte de un plan de mejora de carácter plurianual.
— Debe explicitar los objetivos, los procedimientos y actuaciones previstas, las personas responsables de su ejecución, los recursos y apoyos necesarios, un calendario para su cumplimiento y un plan para su evaluación.
— Debe implicar a las personas, desde una orientación participativa y bajo el impulso asociado a un liderazgo efectivo por parte de la Dirección".

3. PLAN DE EVALUACIÓN DE CENTROS

La metodología de la evaluación vendrá determinada por el objeto de la misma, de los *objetivos y funciones* de la evaluación en cada caso concreto y que vienen *condicionados por el modelo evaluativo* adoptado, con el fin de realizar propuestas de mejora. Estas líneas compartidas por numerosos autores (De Miguel, 1994; De la Orden, 1995; Pérez Juste, 1992, 1995), muestran el hilo conductor de los aspectos principales a tener en cuenta a la hora de realizar el diseño de la evaluación: objetivos y criterios de evaluación, modelos y procedimientos de la evaluación.

3.1. Objetivos y criterios de evaluación

De la Orden (1995) señala que para formular un juicio de valor sobre la organización de un centro docente es necesario compararlo con una instancia de referencia, con lo que esta organización será adecuada o inadecuada comparada con un referencial determinado. La determinación de esta instancia de referencia supone el establecimiento de unos criterios de evaluación. Estos criterios pueden ser múltiples: de adecuación, de deseabilidad de la organización, de bondad, de funcionalidad, de eficacia... Ya hemos visto al principio del capítulo cuando señalábamos las cuestiones que Nevo tenía en cuenta para definir la evaluación la importancia de los objetivos y criterios.

3.2. Modelos de Evaluación

Establecidos los **objetivos y criterios** debemos delimitar el contenido a evaluar, es decir, delimitar los aspectos de los centros educativos que deben ser evalua-

dos. Para ello necesitamos un MODELO TEÓRICO que permita captar las relaciones (de forma global) de todos los factores que intervienen. La aceptación de uno y otro modelo condicionará toda la evaluación: sus objetivos, sus procesos, sus técnicas metodológicas,... tal como hemos visto en el apartado referido a modelos.

3.3. Procedimientos de evaluación

Como hemos estado viendo a lo largo de todas estas páginas es difícil plantear un único proceso metodológico si partimos de la posibilidad de aplicar una gran diversidad de metodologías en función de los objetivos seleccionados.

Todo proceso evaluativo está condicionado por, en primer lugar, el concepto de evaluación del que parta. Tyler (1950) que la consideraba la evaluación en función de los objetivos alcanzados, propone un proceso evaluativo con ocho fases: establecer los objetivos, ordenarlos en clasificaciones amplias, definirlos en términos de comportamiento, establecer situaciones adecuadas para que pueda demostrase la consecución de éstos, explicar los propósitos de la estrategia a las personas responsables, seleccionar o desarrollar las medidas técnicas adecuadas, recopilar datos de trabajo y comparar los datos con los objetivos de comportamiento.

Cronbach (1963), añade el aspecto referido a la toma de decisiones en el proceso. Scriven (1967) integra la validez y el mérito de lo que se realiza.

En los años 60 y 70 los procesos van orientados a la comparación con los resultados de otros sujetos, centros,... lo que hace necesario encontrar indicadores operativos en este sentido.

Nevo (1986) resume en tres fases este proceso:

- Focalizar el problema de evaluación.
- Recoger y analizar datos empíricos.
- Comunicar los resultados a las audiencias.

Casanova (1992) señala las siguientes fases de todo proceso evaluador:

- Definición del plan o diseño de evaluación.
- Recogida sistemática de datos.
- Análisis de los datos recogidos.
- Valoración de la información obtenida.
- Redacción de las conclusiones.
- Toma de decisiones.

De Miguel (1994) hace referencia a la necesidad de tomar decisiones sobre el modelo a utilizar, los instrumentos a través de los cuales vamos a recoger información y determinar los agentes implicados en la evaluación.

De la Orden (1995), señala las siguientes operaciones:

- Plan y diseño de evaluación.
- Desarrollo de instrumentos.
- Recogida de información.
- Análisis e interpretación de los datos.

Posteriormente este mismo autor (De la Orden et al., 1997) señala haciendo referencia al modelo sistémico de Universidad, establecería las siguientes fases:

- Definición constitutiva del modelo de calidad.
- Definición operativa del modelo de calidad.
- Búsqueda de variables de los diferentes sectores del modelo CIPP.
- Evaluación de las variables.
- Selección de los instrumentos de medida.
- Recogida de información.
- Codificación y análisis de datos.
- Evaluación del modelo.
- Validación del modelo.

Pérez Juste (1992) señala que el plan de evaluación incluye la selección de variables, los procedimientos de recogida de información, las fuentes a las que se acudirá para recogerla, el enfoque seleccionado (modelo). Incluso, estas dimensiones deben ir acompañadas de otras de carácter técnico como el asentimiento y la implicación de los que van a ser evaluados, la negociación (Villar Angulo, 1988) y, por supuesto, un sistema de decisiones y seguimiento.

Si nos movemos en programas específicos, por citar alguno simplemente tenemos:

- Programa GRIDS (Eliot et al., 1986) que incluye cinco fases: de preparación, de revisión inicial, de revisión específica, de acción para el desarrollo, de revisión o reiniciación.
- Proyecto IBS/A de Cranston (1991) que recoge las siguientes fases en el diseño y desarrollo del proceso de evaluación de centro educativo:
- Fase de planificación/ negociación/ sensibilización/;
- Fase de autorrevisión; fase de decisión y,
- Fase de seguimiento.

De Miguel (1994) señala cinco fases a utilizar en la metodología de evaluación muy en concordancia con la propuesta anterior: fase de planificación/negociación, aplicación/autorrevisión, valoración, elaboración de programas de mejora y formación y, por último, la fase de seguimiento, revisión e reinicio.

Por último hacer referencia a proceso seguido por el Modelo Europeo de Gestión de la Calidad en donde, con la idea de establecer vínculos entre la autoevaluación, el proceso de planificación y los planes de acción, establece las siguientes fases en la primera de sus modalidades: proceso de autoevaluación y recogida y análisis de datos tradicionales de planificación; definición de acciones de mejora y establecimiento de prioridades (Plan Anual de Mejora), al tiempo que la Dirección elabora el plan de la organización a partir de la planificación estratégica; elaboración de un plan integrado y, por último, implantación y revisión.

A modo de conclusión, la metodología que sigamos en nuestros procesos de evaluación están determinados por los modelos y enfoques teóricos en los que en-

marquemos nuestros objetivos evaluativos, lo que condicionará tanto los instrumentos en la recogida de información con su análisis posterior.

3.4. Técnicas para la recogida de datos

Fijados los objetivos, decidido el modelo así como el procedimiento de la metodología es necesario recabar información, es decir, qué técnicas vamos a utilizar para la recogida y registro de la información.

Gran cantidad de los datos necesarios para evaluar un centro educativo es generado a través de análisis de documentos ya sean personales, públicos, formales, informales…

Solamente las vamos a citar pues éstas están extensamente desarrolladas en multitud de manuales. Lo que hay que decir es que, dado la complejidad de los centros educativos, nos obliga a acudir a fuentes muy diversas de recogida de información independientemente del paradigma en que nos movamos. Ya hemos visto que de la misma forma que es muy difícil encuadrar un determinado modelo dentro de una determinada línea, igualmente resultará muy difícil utilizar sólo técnicas de una determinada metodología para afrontar la evaluación educativa tal y como hoy la entendemos. Así tenemos:

1. ***Los cuestionarios***[18] que pueden ser aplicados por el evaluador o bien convertirse en autoinforme. Esta técnica constituye la más usada en la evaluación institucional y por esta misma razón dos parecen ser los enemigos de la eficacia de los cuestionarios: el uso inadecuado en circunstancias en la que no es necesario y su elaboración poco reflexionada (López Mojarro, 1999). Mediante ellos podemos obtener información de un gran número de personas y cuantitativizar esos resultados.

   Podemos usar cuestionarios que elaboremos nosotros para evaluar algún aspecto concreto o, podemos echar mano de otros ya construidos: por ejemplo, el cuestionario para la Evaluación del Clima del Centro de Halpin y Crof (1963), QUAFE 80 o cuestionario para el análisis del funcionamiento de la Escuela de Darder y López (1985), los cuestionarios desarrollados por Isaacs y col. (1977)[19], cuestionarios para la evaluación de centros de educación infantil-ACEI, cuestionarios desarrollados en el Plan Eva tanto en primaria como en secundaria (1990), cuestionarios de-

[18] La obra presentada por Peralta Sánchez, F. y Sánchez Roda, Mª (1998): *El plan de evaluación: instrumentos*. Escuela Española. Madrid; presenta un amplio repertorio de cuestionarios, escalas de observación, etc. que permiten elaborar un Plan de Evaluación del Proyecto Curricular; del proceso de enseñanza y aprendizaje; del Plan de Acción educativa... como de otros aspectos educativos.

[19] Cuestionarios sobre las tareas típicas del profesor, sobre los objetivos relacionados con la labor educativa, sobre la dirección del centro, encuestas para alumnos de bachillerato y COU sobre el sistema de dirección.

sarrollados por el Modelo Europeo de Gestión de Calidad (1994), cuestionarios desarrollados por de Miguel y otros para la evaluación globalizada de los institutos de educación secundaria.

2. ***Documentos institucionales:*** Proyecto Educativo, Proyecto Curricular, Reglamento de Régimen Interno, Programación General Anual, actas de reuniones...
3. ***La entrevista*** (en sus diferentes formatos: formal -cerrado, semi-abierto, cerrado- e informal). Se la define como una conversación intencional con dos objetivos principales: la investigación y la orientación. Por esta razón sus funciones son de carácter diagnóstico, evaluador, investigador, orientador y terapéutico.

   A través de ella se pretende alcanzar una máxima comprensión de los fenómenos, observando opiniones subjetivas sobre algún aspecto del objeto evaluado y matizar aquellas cuestiones de las que tengamos dudas o bien para prefijar la estructura del cuestionario. Pero también tiene sus inconvenientes puesto que precisa de mucho tiempo y dedicación en la preparación y aplicación y puede darse un sesgo en el entrevistador. Entre las características que señala Casanova (1992) para que tenga provecho en un proyecto de evaluación tenemos: definición clara de sus objetivos, delimitación de la información que se desea obtener, conocimiento del tipo de relación que mantienen entrevistador y entrevistado, garantía para el entrevistado de la reserva y confidencialidad de la información y registro de la conversación mantenida.
4. ***Listas de control*:** consiste en una tabla de doble entrada en la que recogemos información sobre aquellos datos observables de lo que queremos evaluar. En una entrada se reflejan los rasgos que queremos registrar y, la otra si se dan o no. Permiten guiar y disciplinar la atención del observador en aquellos aspectos que considera relevantes.
5. ***La observación*** en sus diferentes dimensiones: en formato abierto (el observador toma nota de las reacciones o conductas de los sujetos); autorregistro (el observador registra sus propias reacciones o conductas); sistemas de signos (se cuentan las conductas específicas cada vez que ocurren); sistemas de categorías (la conducta observada es clasificada en categorías). Entendiendo por observación, junto con Casanova (1992:75), "un examen atento que un sujeto realiza sobre otro u otros sujetos sobre determinados objetos y hechos, para llegar a un conocimiento mayor sobre ellos mediante la obtención de una serie de datos, generalmente inalcanzables por otros medios". Lo importante es que se tenga claro cuál va a ser el objeto de la observación, el responsable, la fecha de finalización de la observación y los aspectos a observar y su registro mediante instrumentos necesarios.

La observación podrá ser participante en la medida que el observador se integra al grupo que quiere observar, o no participante, en la medida que se mantiene al margen de las actividades y relaciones del grupo.

6. ***Encuesta:*** obtención de la información a través de la aplicación de cuestionarios orales o escritos con el fin de averiguar y describir las condiciones existentes en el desarrollo de la situación que evalúa, describir normas de conducta o funcionamiento, determinar relaciones existentes. Para ello deberemos conocer la finalidad que persigue la aplicación de la encuesta, la población de destino y los recursos de los que se dispone.
7. ***Las escalas de estimación*** nos permiten la valoración/ordenación de los fenómenos, en la que cada elemento observado se valora con la gradación de categorías que consideremos más oportuna en función de que la escala sea numérica (de 1 a 5, p.e.), gráfica o descriptivas (muy similares en la que se sustituyen los números por descripciones concisas (muy importante a nada importante, p.e.). Así como el cuestionario suele ser cumplimentado por aquellos que no participaron en su elaboración, las escalas suelen ser cubiertas por el propio evaluador y, en ciertos casos, como pueden ser en procesos de autoevaluación, por los sujetos evaluados.

   Podríamos señalar algunos ejemplos como la Escala de evaluación de un centro docente de Enseñanza Básica (CEDODEP, 1996), Escala de Clima Social de Moos y Triket (1974) o la de Estimación del Clima social de Medina Rivilla (1988).
8. ***Inventarios*** como el desarrollado por Villa Angulo y Carlos Marcelo (1991) sobre el ambiente de clases universitarias.
9. ***Visitas in situ***, documentos: proyectos educativos, actas, programaciones, etc., revisiones de expertos, paneles.
10. ***Grupos de discusión:*** son similares a la entrevista puesto que lo que se busca es obtener información sobre un determinado tema a través de un debate en el que cada uno expone su perspectiva. La diferencia con la entrevista estriba en que aquí no existe la figura del entrevistador, con lo que el grupo se monta entre iguales. A través de esta técnica podemos detectar actitudes personales, intereses, etc. Como la actuación del entrevistador o coordinador es nula, se centrará solamente en que los demás se expresen y coordinar, es importante la grabación de las sesiones para su posterior transcripción análisis.
11. ***Triangulación***: aunque la podemos considerar como una técnica de análisis de datos la incluimos aquí por ser también una técnica propia de la investigación social que surge por la necesidad de asegurar la veracidad de lo observado y de su fiabilidad. Se trata de multiplicar las observaciones con el fin de lograr una mayor validez en la investigación, buscando

la coincidencia plena de las observaciones (se recomiendan tres, aunque depende del investigador). Esta triple observación puede ser realizada desde tres fuentes de información (triangulación de fuentes: profesores, padres y alumnos, p.e.), desde tres métodos y técnicas de observación (entrevistas, cuestionarios y observación), en momentos y circunstancias diferentes (triangulación temporal), en lugares diferentes del centro (triangulación espacial), o por evaluadores distintos (profesores, inspectores, alumnos,...). Es una técnica de corte más cualitativo, por lo que suele usarse cuando recogemos información desde técnicas de observación o entrevista.

3.4.1. *Algunas aportaciones en la evaluación de centros*

Son numerosas las aportaciones respecto a los instrumentos de recogida y registro de información en evaluación de centros. Algunos de ellos quedan reflejados en las obras de Pérez Juste (1992) o Cardona (1994). Nosotros vamos a señalar también alguno de ellos que consideramos de un gran interés por dos motivos principalmente: el primero de ellos obedece, al igual que nos dice el autor, a un fin de divulgación; el otro, a ir asentando el "QUÉ EVALUAR". Cada uno de estos instrumentos o guías para la evaluación nos van a ir ofreciendo una serie de dimensiones, subdimensiones, aspectos, que son evaluables en el centro educativo y que nos permitirán ir haciendo un mapa de lo que se está evaluando en nuestros centros educativos.

Pauta para la evaluación de un centro educativo de García Hoz (1975). Esta pauta que se estructura en cinco dimensiones, está integrada por 100 s:

— Alumnos (15 s).
— Personal Educador (15 s). Se incide en dimensiones humanas, formación técnica de profesores y directivos y programación de actividades.
— Medios materiales (10 s).
— Actividades internas (40 s): se analizarían funciones del profesorado.
— Relaciones centro-familia-sociedad (20 items). Se analiza la existencia de comunidad educativa.
— El peso en esta pauta recae, en contraposición al anterior en el análisis de las funciones del profesor, del rigor de lo programado y su logro.

Aportación de D. Isaacs (1977). Lo que pretende el autor es evaluar a los centros a través de diferentes cuestionarios, desde la perspectiva de la opinión de los profesores y directivos, con una orientación a la detección de necesidades. Esta perspectiva, incluye en nuestra opinión algunos aspectos metodológicos que serán tenidos en cuenta posteriormente en otros modelos e instrumentos de evaluación de centros, al considerar quién evalúa (profesores, equipo directivo) y su fin (orientado a la mejora a través de la detección de necesidades).

El autor presenta tres cuestionarios centrados en las tareas del profesor (orientación a estudiantes, perfeccionamiento continuo, normativa de convivencia), objetivos educativos (objetivos relacionados con actitudes referidos a orden, atención, responsabilidad... y objetivos relacionados con actividades como participación, cuidado del material, técnicas de estudio...) y dirección del centro (dirigido al centro educativo, al director, a los directivos medios y profesores y padres de familia y entorno).

Guía para la evaluación de Instituciones de Nivel Superior de Escotet (1984). En esta guía se recogen multitud de preguntas relacionadas con las siguientes dimensiones evaluativas:

— Organización y estructura de la Universidad en donde se analiza la estructura y organización de la Universidad a través de 49 s.
— Estudiantes. A través de más de 300 preguntas analiza otras subdimensiones relativas a las causas de deserción estudiantil, los programas de actividad estudiantil o los aspectos que determinan el rendimiento académico.
— El currículum. A través de 130 preguntas se recoge información sobre los objetivos, organización del currículum, actividades, recursos y materiales, sistema de evaluación, departamentalización curricular, sistema de evaluación.
— Planta física (120 preguntas) analizan la relación entre espacios y su utilización, lugar, edificios, servicios, salón de la clase, oficinas, mantenimiento.
— Servicios de extensión e investigación (80 preguntas). Se analizan aspectos relativos a los objetivos, organización, servicios y procedimientos.
— Recursos humanos.

El mayor peso se centra en las labores de los estudiantes y el currículum, frente a los servicios de extensión e investigación.

Quafe-80 o Cuestionario para el análisis del Funcionamiento de la escuela. Es un instrumento que fue creado para favorecer la evaluación interna en la escuela. Se distribuye en dos grandes dimensiones. El primero referido al proyecto educativo, que consta de doce items agrupados en torno a dos subdimensiones referidas al diseño del proyecto y su grado de implantación y, el segundo factor, que hace referencia a la estructura y funcionamiento del centro: consta de 14 items agrupados en las subdimensiones de gestión, actuación personal, recursos y relaciones.

La estructura de cada ítem es igual para las dos dimensiones: definición del aspecto a evaluar y descripción de cinco situaciones en las que puede hallarse en el centro con referencia a dicho ítem.

Plan EVA (1992) del MEC. Este plan se intenta para ayudar a los centros docentes a superar sus dificultades que van a encontrar para acometer sus propios

procesos de evaluación interna, ofreciendo a los centros un modelo y un conjunto de indicadores de evaluación que facilite la autoevaluación de su organización y funcionamiento.

Entre otro de los objetivos que persigue (Plan EVA, 1996) es mejorar la organización y funcionamiento de los centros en los que se lleve a cabo la evaluación, estimulando y apoyando la consecución del proceso de evaluación iniciado en cada uno de ellos.

Entre los objetivos complementarios que persigue se encuentra, muy resumidamente, el de difundir una cultura evaluadora entre todos los agentes educativos, continuar ensayando los métodos, las técnicas y los procedimientos de análisis para la evaluación de los centros, así como la de contribuir a la actualización de los Equipos de inspectores en actividades evaluadoras de los centros.

Se lleva a cabo a través de un proceso de evaluación formativa externa dirigido a un grupo de inspectores y con la audiencia de todos los agentes y sectores implicados en el centro.

Su objeto es identificarse con el modelo sistémico de evaluación de centros docentes adoptando metodologías mixtas que integran enfoques cualitativos y cuantitativos.

Los instrumentos en los que se basa son los cuestionarios, entrevistas, etc. Centrándose en aspectos relacionados con la entrada y proceso, es decir, con la organización y funcionamiento de los centros; si bien, tiene en cuenta también aspectos de contexto y producto.

Estos aspectos han sido agrupados en cinco grandes dimensiones:

— Elementos contextuales y personales.
— Proyectos (permite obtener datos sobre la valoración de estrategias de programación en los diferentes proyectos: PDC, PCE y PA).
— Organización y funcionamiento (obtiene información sobre los órganos de gobierno, equipos de ciclo y Departamentos y sobre la comunidad educativa).
— Procesos didácticos. Proporcionan información sobre aspectos referidos a la relación del profesorado en el aula y tutoría, sobre los aspectos metodológicos y evaluación.
— Resultados: proporciona información sobre los logros obtenidos teniendo en cuenta el rendimiento académico, las actitudes de los alumnos y la satisfacción de los distintos sectores de la comunidad educativa.

Aportación de De Miguel (1994) para la Evaluación de los Institutos de Educación secundaria. Partiendo del modelo CIPP, este autor pretende que sean los propios agentes educativos los que evalúen continuamente su trabajo. De este modo proponen una evaluación que suponga la reflexión crítica, contextualizada, que implique una visión holística y multidimensional del centro en la que participen todos los agentes educativos, favoreciendo la triangulación. Por tanto se trata

de una evaluación formativa, integrando tratamientos cualitativos y cuantitativos, orientada al cambio y que permita tanto la evaluación interna como externa.

Este autor, elabora 15 cuestionarios que permiten recoger información sobre seis dimensiones del centro encardinadas dentro del modelo CIPP:

— ***Contexto:*** incluiría las subdimensiones entorno sociocultural, entorno familiar, características del centro y hábitat.
— ***Entrada:*** con las dimensiones referida infraestructura (instalaciones y recursos, servicios, personal y apoyos externos).
— ***Proceso.***
— ***Estructura:*** órganos de gobierno, equipos de profesores, responsables de la coordinación didáctica, responsables de servicios y órganos de participación de la Comunidad Escolar.
— ***Función:*** programas educativos, de funcionamiento interno y de administración y servicios.
— ***Clima institucional:*** modelo organizativo, nivel de comunicación, participación...
— ***Producto (resultados):*** desarrollo organizacional del centro, calidad de los aprendizajes y desarrollo humano, desarrollo profesional y satisfacción y reputación.

Plan ACE. Plan Gallego de Evaluación de Centros educativos (1995). A partir de la Ley Orgánica 9/1995, del 29 de noviembre, sobre la participación, la evaluación y el gobierno de los centros docentes (BOE del 21 de noviembre de 1995). La Consellería de Educación e Ordenación Universitaria pone en marcha en el curso 1995/1996 el plan de Evaluación de centros (Plan ACE).

Inicialmente este Plan es enfocado como un plan de evaluación externa de los centros, pero los documentos pueden ser utilizados para una evaluación interna de los mismos; ya que, dicho plan, persigue la mejora de la Organización y funcionamiento de los centros como favorecedora de la calidad de la enseñanza a través de un proceso continuo de evaluación interna de los propios centros.

Este Plan se constituye como una clara orientación a convertirse como plan de evaluación interna, asumido y desenvuelto por todos y cada uno de los centros de la Comunidad Autónoma de Galicia.

La finalidad del plan, tal como pone de manifiesto el documento elaborado por la Consellería de Educación y Ordenación Universitaria es la de favorecer, como poníamos de manifiesto, la calidad de la enseñanza a través de la mejora de la organización y del funcionamiento de los diferentes centros docentes. Este objetivo general, es complementado por otros más como son la de difundir una cultura evaluadora entre todos los centros educativos así como la de aportar a los centros técnicas e instrumentos que faciliten la articulación de su propia autoevaluación. Objetivos que están en concordancia con los señalados anteriormente en el Plan EVA.

La evaluación, planteada en un principio como de carácter cualitativo, se centra en la actividad educativa implicando a todos los agentes, como la administración educativa y sus servicios, las instituciones locales y la comunidad educativa (padres, alumnos y personal docente).

Dicha evaluación utiliza diversas técnicas de recogida de datos sobre diferentes aspectos de la vida del centro desde distintos ángulos y perspectivas, tales como la entrevista, reuniones, cuestionarios y análisis documental. La finalidad de estos instrumentos, únicamente, la de sintetizar el proceso de recogida de datos.

La información recogida por las anteriores técnicas, será analizada por un equipo evaluador (inspección si se trata de evaluación externa y, equipo directivo si fuese interna), el Consejo Escolar y el Claustro de Profesores, teniendo un carácter eminentemente formativo.

Este plan comprende los cuatro ámbitos del modelo CIPP, es decir, contexto, entrada, proceso, producto, siendo las dimensiones a evaluar las mismas que las expuestas en el Plan EVA.

Modelo Europeo de Gestión de la Calidad. MEC (1998). La adopción del modelo de la EFQM por el Ministerio de educación y cultura se basa, sobre todo, por su potencial de adecuación a los entornos educativos públicos. Este modelo combina el interés por las personas con los procesos y resultados. El Ministerio elabora una "Guía para la autoevaluación" que viene acompañada por otras cinco publicaciones que complementan y configuran un conjunto de materiales dirigidos a los centros educativos con el fin de proporcionar las herramientas necesarias a todos los centros que quieran aplicar el modelo: "Guía para la autoevaluación", "cuestionario para la autoevaluación", "formulario para la autoevaluación", "caso práctico de autoevaluación de un centro -exposición del caso práctico de autoevaluación de un centro - valoración del caso-" y "Plan Anual de Mejora. Materiales para el Diagnóstico".

La estructura permite al centro aprender mediante la comparación consigo mismo y le ayuda a la planificación, en la definición de estrategias, en el seguimiento de los progresos seguidos y la corrección de errores como efecto de su actividad.

ACTIVIDADES PARA EL ALUMNO

1. Realiza un esquema comparativo de los modelos señalados por cada una de las clasificaciones señaladas. Señalando los aspectos parecidos y las diferencias esenciales en cada uno de estos modelos (acude a la bibliografía reseñada).
2. Elabora una lista de las dimensiones, subdimensiones, y aspectos que evaluarías en un centro educativo. Señala en cada caso el instrumento/s o técnicas que utilizarías para obtener la información, así como a quién se lo dirigirías.

BIBLIOGRAFIA COMENTADA

ÁLVAREZ FERNÁNDEZ, M. (1998). *El liderazgo de la calidad total.* Madrid: Escuela española.

En esta obra se ofrecen argumentos fundamentados con el fin de provocar una reflexión sobre la necesidad e importancia del ejercicio de un liderazgo de carácter integrador en el centro docente. El trabajo analiza la gestión de la calidad como reto educativo; las características del líder y los pilares del liderazgo; la visión del profesor deseable… Además de ofrecer multitud de sugerencias, técnicas e instrumentos para los formadores.

CARDONA ANDÚJAR, J. (Coord.) (1994). *Metodología innovadora de evaluación de centros educativos.* Madrid: Sanz y Torres.

Con un estilo muy didáctico el autor desarrolla diferentes aspectos a tener en cuenta en la evaluación de los centros educativos: concepto, objetivos, variables, enfoques, modelos,… ofreciendo innumerables referencias bibliográficas sobre el tema, orientaciones para el estudio, esquemas, etc…

DE MIGUEL, M. (1989). "Modelos de investigación sobre organizaciones educativas". *Revista de Investigación Educativa,* 7, 13, 21-56.

Este artículo es de especial relevancia por la importancia que ha tenido en posteriores publicaciones a la hora de clasificar los modelos de evaluación de centros siendo guía para numerosos autores.

DE MIGUEL, M.; MADRID, V.; NORIEGA, J. Y RODRÍGUEZ, B. (1994). *Evaluación para la calidad de los Institutos de Educación Secundaria.* Madrid: Escuela Española.

Este texto puede ser muy útil como guía concreta para las instituciones educativas que quieran poner en marcha procesos de autoevaluación o autorrevisión, al ofrecer propuestas e instrumentos reales como estrategia para mejorar la calidad de los centros educativos.

LÓPEZ MOJARRO, M. (1999). *A la calidad por la evaluación.* Madrid: Escuela Española.

Es una obra muy interesante que afronta la calidad desde una intención clara y equilibrada. Ofrece un modelo de evaluación sobre la calidad total en el que sustentar los planes de mejora. Con innumerables ejemplos, tablas y gráficos, el autor logra que el lector obtenga una visión práctica del proceso de evaluación desde este modelo.

STUFFLEBEAM, D.L. Y SHIKFIELD (1987). *Evaluación sistemática. Guía teórica y práctica.* Barcelona: Paidós.

Obra de lectura obligada para aquellas personas interesadas en la investigación evaluativa o para aquellos que se están iniciando. Este libro realiza, después de dar una visión sobre el estado de la evaluación así como del análisis de métodos alternativos de evaluación, un repaso sobre ocho modelos referidos al trabajo evaluativo de las más importantes figuras de la investigación evaluativa.

FUENTES DOCUMENTALES

1. Fundación Europea para la Gestión de Calidad.
EUROPEAN FOUNDATION FOR QUALITY MANAGEMENT
Aveneu des Pléiades, 19
1200 BRUSEELS (Bélgica)

2. Plan Nacional de Evaluación.
Ministerio de Educación y Ciencia.
MADRID.
www.mec.es

3. National Association for Secundary School Principals:
1904- Association Drive
RESTON, V.A. 22091 (USA)

CAPÍTULO VI

EVALUACIÓN INSTITUCIONAL Y MEJORA DE LA CALIDAD EN LA ENSEÑANZA SUPERIOR

Leonor Buendía Eisman[1]
Beatriz García Lupión[2]
Universidad de Granada

OBJETIVOS

1. Conocer la situación actual de la universidad y de la educación superior
2. Desarrollar cómo ha evolucionado la evaluación institucional hacia la mejora de la calidad de la enseñanza superior
3. Conocer cómo surge la evaluación institucional en Europa
4. Conocer los distintos programas de evaluación institucional en España
5. Desarrollar los distintos modelos de evaluación institucional, así como su puesta en marcha en distintos países del mundo

CONTENIDOS

Introducción
1. Aproximación teórica a la universidad de nuestros días
2. Educación superior en la universidad actual
3. Evaluación institucional y mejora de la calidad de la enseñanza
4. Evaluación institucional en el contexto europeo
5. Evaluación institucional en España
   5.1. Programa Experimental de Evaluación de Calidad del Sistema Universitario
   5.2. Proyecto Piloto Europeo de Evaluación de la Calidad de la Enseñanza Superior
   5.3. El Plan Nacional de Evaluación de la Calidad de las Universidades
6. Modelos de evaluación aplicados a la institución universitaria

Actividades para el alumno
Bibliografía comentada

[1] Facultad de Ciencias de la Educación. Área de Métodos de Investigación y Diagnóstico en Educación. Universidad de Granada. Campus universitario de Cartuja. Tfno: 958-243755/958-243762. E-mail: lbuendia@platon.ugr.es
[2] Facultad de Ciencias de la Educación. Área de Métodos de Investigación y Diagnóstico en Educación. Universidad de Granada. Campus universitario de Cartuja. Tfno: 958-243755/958-243762. E-mail: bglupion@platon.ugr.es

INTRODUCCIÓN

En los últimos años, la evaluación institucional universitaria se ha convertido en un tema prioritario en toda Europa. Por un lado, la sociedad exige cada vez mayor calidad de los servicios que prestan las universidades, por otro, el gobierno debido a la mayor autonomía de las universidades, le exige a éstas una rendición de cuentas.

El número de personas que acceden a la universidad "ha tocado techo", por lo que las demandas han pasado de centrarse en la cantidad para mirar a la calidad de dicha institución, como objetivo prioritario en los próximos años. Parece razonable que, si en la época de rápido crecimiento del sistema, las prioridades se centraron en dar cabida al mayor número de estudiantes que solicitaban cursar estudios en la universidad, ahora y cumplido parcialmente el objetivo de cantidad las prioridades se dirijan a cubrir el objetivo de calidad.

La sociedad, por otra parte, exige titulados formados, capaces de hacer frente a las demandas de una sociedad cambiante. Desde un modelo de universidad cerrado y con contenidos universales, nos encontramos ante una nueva situación en la que los cambios tecnológicos, el aprender a aprender como forma de reciclaje y adaptación, la multidisciplinariedad y sobre todo, la contextualización de los aprendizajes, exigen una universidad abierta a la sociedad y competitiva, o sea, una continua revisión de la calidad de lo que ofrece, para no quedar obsoleta y al margen de las demandas de los nuevos tiempos.

A su vez, las universidades, al tener mayor autonomía, tienen que rendir cuentas a la sociedad de lo que hacen con el dinero que reciben de los contribuyentes para formar a los estudiantes, y es que el contribuyente, cubre una parte fundamental de los costes en todas las universidades europeas. Debido al elevado nivel de financiación pública y a que la autonomía no implica necesariamente un mejor rendimiento de la institución, se hace necesario evaluar dicho rendimiento y dar a conocer los resultados.

Esta situación, ha llevado a la mayoría de los países desarrollados, a poner en marcha sistemas de evaluación institucional. España es uno de los países europeos que en la actualidad está evaluando sus universidades, a través del Plan Nacional de Evaluación de la Calidad.

La evaluación en este contexto, no es un fin en sí misma. Tiene sentido en la medida en que contribuye a mejorar la calidad, aunque no es suficiente para ello, ya que una buena calidad supone entre otras cosas, recursos suficientes y mecanismos apropiados de gestión. Pero esto exige una evaluación: si no sabemos si algo está bien o mal, en qué aspecto o medida, no podremos orientar el esfuerzo institucional ni desarrollar estrategias de superación (Martínez, 2000). Por tanto, la evaluación no se utiliza sólo para verificar el funcionamiento de la institución universitaria y los resultados obtenidos, sino que se concibe como un método para mejorarla. El

proceso que se está siguiendo en la mayoría de los países europeos, combina la autoevaluación y la evaluación externa, con el objetivo de detectar las fortalezas y debilidades de la institución y, llevar a cabo acciones de mejora.

En este capítulo, se pretende dar a conocer la situación actual de la universidad. Se presentan los distintos programas en los que ha participado España, así como las principales características de los modelos de evaluación institucional que mayoritariamente se han utilizado en distintos países del mundo.

1. APROXIMACIÓN TEÓRICA A LA UNIVERSIDAD DE NUESTROS DÍAS

La universidad actualmente tiene ante sí el reto de adaptarse a los grandes cambios ocurridos en las últimas décadas en la sociedad. Esta labor será necesaria para que siga estando a la cabeza en la formación de los jóvenes y en la creación de nuevos conocimientos, tal y como viene haciendo desde hace siglos.

Los problemas y dilemas a los que se enfrenta la universidad son: los cambios tecnológicos de la llamada sociedad de la información, la creciente demanda tanto de mayor calidad como de reciclaje continuado, el empleo de nuevos métodos de formación, la necesidad de aunar la especialización con el cultivo de las ciencias multidisciplinares, la autonomía universitaria, etc.

Las cuestiones que le surgen a la universidad de nuestros días son relativas a cómo hacer compatible la dimensión social de servicio público con las exigencias de calidad que son propias de una institución dedicada al desarrollo y transmisión de la ciencia, la tecnología y la cultura superior (Quintanilla, 1996).

Es en el nuevo contexto donde la universidad tiene la oportunidad de demostrar su capacidad de adaptación y el liderazgo social e intelectual, que debe mantener. Para ello se deben adoptar cambios drásticos tanto en la forma como en el diseño de su organización interna y en las relaciones con la sociedad. Los grandes problemas a los que esta institución tendrá que hacer frente en el futuro provienen de dichas relaciones. Consecuencia de esto será el nuevo esquema de relaciones entre universidad y resto de la sociedad.

Una universidad de espaldas a los cambios sociales y tecnológicos supone un despilfarro económico que ningún país se puede permitir. Estas relaciones exigen a la universidad, "enseñar para aprender" en una nueva época en la que los conocimientos de hoy quedan obsoletos mañana y que por lo tanto la apertura, diversificación, necesidad de reconversión y adaptabilidad a los tiempos, deben ser características de una formación universitaria, en los momentos actuales.

Una parte importante de la actividad universitaria, está estableciendo una vinculación diferente con la sociedad. Esta parte tiende a comportarse, a largo plazo, como un sector interno en el sistema de producción de un país, unido al resto por un conjunto de inputs (medios de comunicación y de información, laboratorios

de ensayos y de investigación, etc.) y de outputs (investigación aplicada, formación, particularmente bajo el concepto de "vocational training", etc.) (Bricall, 1998).

Nuevos problemas surgen en el día a día de la vida de las universidades. Dice Quintanilla (1996), que son "las reivindicaciones locales de creación de nuevas universidades, tan frecuentes en los últimos años, o las decisiones de los poderes públicos de cambiar, crear, suprimir y recombinar centros universitarios como si se tratara de colegios de barrio, síntomas de las nuevas situaciones". Estas transformaciones son un producto de los países desarrollados y tienen consecuencias importantes tales como la nueva conceptualización de la "Universidad de masas", concepto nacido para dar respuesta a las necesidades de un sistema productivo y al derecho a la educación de todos los ciudadanos en la sociedad actual.

Lo que caracteriza a la denominada universidad de masas es que gran parte de la población accede a los estudios universitarios, existiendo demandas sociales más amplias y exigentes.

Las universidades tradicionales no estaban pensadas para una gran demanda social, existiendo un desajuste con las tendencias de la realidad actual.

Ante esta situación las universidades, según Rojo (1993) tienen que aprender a afrontar este crecimiento sin un deterioro de la docencia ni de la formación profesional e investigadora. La autonomía y la independencia son esenciales a la hora de estimular la buena docencia, formación e investigación de las universidades públicas. Esta autonomía va unida a la autogestión y autofinanciación, para entenderla en su máxima expresión, pero si no se da, entonces se propondrán planes de evaluación del rendimiento y de resultados, ya que las inversiones tienen que estar ligadas a la eficiencia y rendimiento social. Este proceso de evaluación será dinámico y cuyo fin será la mejora de la institución en su globalidad.

2. EDUCACIÓN SUPERIOR EN LA UNIVERSIDAD ACTUAL

Blat (1994) hace una reflexión sobre el estado del sistema educativo en niveles superiores:

> *"Si hoy despertara una persona que hubiese estado dormida durante cien años se asombraría de los cambios que ha sufrido la sociedad; sin embargo, el proceso de formación lo encontraría casi igual".*

El reto de la educación superior será aparecer como sistema capaz de responder a las exigencias de la sociedad, traduciéndose éstas en; mayor demanda de educación superior, variación de la demanda de cualificaciones profesionales, elevar los niveles de calidad y eficiencia, etc.

Pero quizás el problema se presenta, tal y como afirma Guerci (1993) *"cuando se piensa en que es posible establecer cortes en la realidad dando como resultado: una educación ajena a los problemas de su tiempo, estéril a la hora de proponer alternativas de vida relevantes, porque no puede establecer un diálogo entre sus construcciones teóricas y las necesidades más elementales del hombre en su cotidianeidad"*.

En opinión de Rojo (1993), existen tres razones en la sociedad moderna que deben ser satisfechas por un sistema de educación superior:

1) Proporcionar un alto nivel cultural y preparación científica a mayores segmentos de la sociedad, lo cual incluye aportar un elemento de formación profesional a aquellos que en la actualidad no progresan en sus trayectorias educativas.
2) Proporcionar una formación eficaz a los profesionales que conformarán la futura espina dorsal de la sociedad.
3) Llevar a cabo investigación que contribuya no sólo al progreso de las fronteras del conocimiento, sino también al desarrollo industrial.

No es fácil combinar estas tres tareas si se intenta al mismo tiempo mantener la calidad y, por otra parte, es dudoso que las universidades sean la mejor elección a la hora de desarrollarlas, teniendo en cuenta las consideraciones planteadas en el apartado anterior.

Resulta, de esta situación de cambio, una universidad para satisfacer las demandas de la sociedad industrial. Será precisamente en esta perspectiva donde se posiciona el memorándum sobre educación superior en la Comunidad Europea y de la Conferencia Permanente de Rectores, Presidentes y Vicerrectores de las Universidades Europeas, que han puesto de relieve su preocupación para que la educación superior en las universidades se trate casi exclusivamente en términos de proporcionar servicios a la comunidad económica.

Remitiéndonos a Van der Molen (1999) argumenta que determinados valores asociados a las universidades, a lo largo de los siglos -como los de juicio independiente, creatividad, y las dimensiones culturales y éticas- son probablemente de igual importancia, si no mayor, y que deben estar basados en el conocimiento científico. Por lo que la educación superior ha de ser entendida tanto como sistema desde el que resolver problemas y desarrollar actitudes, como preparación para el aprendizaje, más que acumular conocimientos para trabajos específicos.

Así Van der Moler, (1999) propone como actividades esenciales en la educación superior: la creación, la transmisión y la aplicación del conocimiento.

Puede considerarse válida esta propuesta, pero no hay que dejar de lado que los sistemas de educación superior de los distintos países difieren en tantas cosas que no existe una única solución que, en un futuro próximo, pueda producir un modelo uniforme generalmente aceptado. La mayoría de las sociedades y de los gobiernos valoran la educación superior como algo beneficioso para la sociedad en

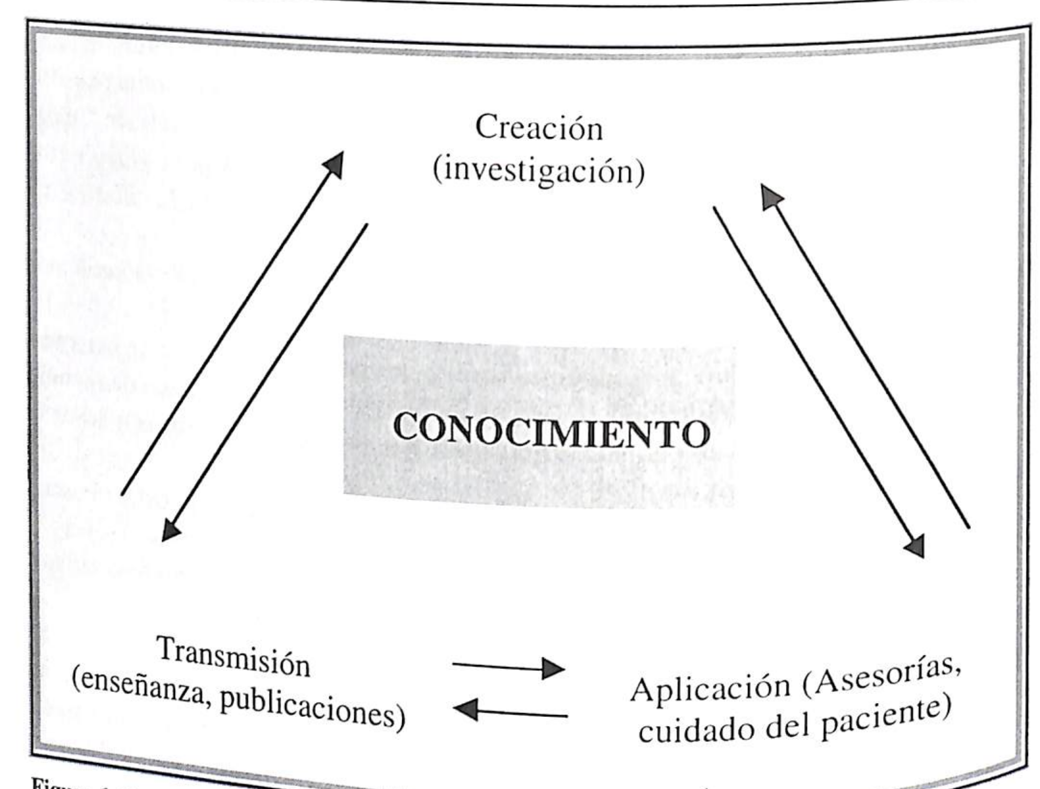


Figura 1. Principales actividades de la educación superior (Van der Moler, 1999).

su conjunto así como para el individuo. Es el éxito y el gran interés que suscita la educación superior lo que ha provocado tensiones organizativas y financieras dentro de la misma, así como en los gobiernos y en los estudiantes. La información disponible debería permitirnos, sin embargo, diseñar arreglos factibles, dado que éstos cuentan con apoyo político y social.

3. EVALUACIÓN INSTITUCIONAL Y MEJORA DE LA CALIDAD DE LA ENSEÑANZA

En los últimos años, la evaluación de la calidad de las instituciones universitarias se ha convertido en un tema prioritario en Europa. Países como Holanda, Francia, Reino Unido y Dinamarca poseen sistemas consolidados de Evaluación Institucional Universitaria. La universidad española no ha quedado al margen de esta corriente y se encuentra inmersa en un amplio proceso de evaluación con el objetivo de mejorar la calidad de nuestras universidades.

Esta predisposición hacia la evaluación de la calidad de las instituciones universitarias, lleva consigo un cambio, se ha pasado de una visión burocrática del sistema a una visión más autónoma. Esta mayor autonomía implica una mayor y

mejor evaluación de las instituciones y una perspectiva diferente que supone, además de una verificación del funcionamiento del sistema, un método para mejorarlo (Hüfner y Rau, 1987).

Ante esta situación, los gobiernos exigen una rendición de cuentas a las universidades, con el fin de demostrar que tanto las decisiones tomadas, como las acciones ejecutadas están dirigidas a la mejora de la calidad (Rodríguez, 1997). Todo esto unido a la exigencia de calidad de los ciudadanos en todos los procesos productivos y en los servicios, la competitividad económica, y la relativa estabilidad tanto financiera como de alumnado, parece indicar que ha llegado el momento de preocuparse por la mejora de la calidad de las instituciones universitarias.

De esta manera, observamos cómo a lo largo de las últimas décadas los gobiernos de los países desarrollados han ido preocupándose por alcanzar distintas metas para la mejora de la calidad de las instituciones universitarias. Así, mientras que en la década de los sesenta, el objetivo prioritario era dar cabida al mayor número de alumnos/as que solicitaban cursar estudios en la universidad, en la década de los setenta, como consecuencia del gasto público en las universidades y la mayor democratización en el acceso a los niveles educativos superiores, la preocupación fue instaurar un sistema de gestión del proceso universitario, que valorara la efectividad y la eficiencia de las instituciones universitarias (Mora, 1998). Pero es realmente en la década de los ochenta, cuando las prioridades se dirigen a mejorar la calidad del servicio que prestan nuestras instituciones universitarias.

Aunque ha sido aceptado por diversos autores el concepto de calidad que postula por un lado, el ajuste a las necesidades sociales, además del cumplimiento de las normas institucionales y la efectividad en el logro de las metas establecidas, debemos tener presente que siguen existiendo serias dificultades para definir la calidad.

Álvarez (1998) dice a este respecto que hoy, la competencia es tan enorme que el producto se va a convertir en oferta, es decir, va a ser sobre todo y ante todo una respuesta a la necesidad y expectativa del cliente considerado no sólo como consumidor, sino como usuario satisfecho del servicio. Con lo que la calidad se convierte en el resultado de un esmerado cuidado de los procesos, de buenas materias primas, del trabajo bien planificado y de la manera en que el producto o servicio se entrega al cliente.

Otra de las características sobresalientes de la calidad va a ser también su complejidad para ser evaluada. Sin embargo, esto no ha sido un impedimento para que investigadores de diferentes países, hayan tratado de evaluarla.

Los enfoques utilizados para valorar la calidad de una institución universitaria, pueden encuadrarse en las categorías que proponen George (1982) y Astin (1985):

La calidad como prestigio. Hace referencia a la reputación académica y social de algunas instituciones, distinguiéndolas de otras. Desde este enfoque la cali-

dad de una institución universitaria se valora a través de juicios de expertos, por lo que no puede ser evaluada objetivamente. La crítica que se le hace a este enfoque radica en su subjetividad.

Calidad como disponibilidad de recursos. Desde este enfoque la calidad de las instituciones universitarias se valora en función de la disponibilidad de los recursos humanos, económicos y físicos. La crítica que se le hace a este enfoque es que no siempre coincide que una institución grande, por contar con mayores recursos, sea de más calidad.

Calidad a través de los resultados. Una institución universitaria es exitosa por el éxito de sus egresados. La mayor crítica a este enfoque es que el éxito de los egresados puede deberse más al contexto o al propio estudiante que a la institución.

La calidad por el contenido. Desde este enfoque la calidad de las instituciones universitarias es valorada en función de los conceptos implicados en el proceso de aprendizaje (currículum, sistema pedagógico, etc.). El problema radica en la dificultad de evaluar dichos conceptos.

La calidad por el "valor añadido". La institución consigue un valor añadido. Para los defensores de este enfoque, las instituciones de más calidad son las de mayor impacto sobre los conocimientos, por ejemplo el cambio conductual o el desarrollo de su carrera. Pero, ¿cómo aislar la influencia del Centro de otras fuerzas sociales?

En la actualidad, como señala Mora (1991: 35) *"no se ha resuelto el problema del enfoque óptimo de la calidad, la búsqueda de la calidad a través de medidas como las indicadas tiene poco que ver con la mejora del sistema como un todo y con la calidad del programa educativo. Sin embargo, todas ellas juntas, debidamente ponderadas, sí deben dar una cierta idea de la calidad de una institución".*

En cualquier caso, afirma Tejedor (1997: 413): *"Parece obvio que una institución universitaria sólo puede alcanzar un razonable nivel de calidad cuando los elementos humanos, financieros y físicos, la enseñanza y la investigación, la organización y la dirección, sean los apropiados para los fines que la institución persigue".*

De alguna manera todos participan en considerar, una institución universitaria de calidad cuando es pertinente, eficiente y eficaz.

La pertinencia podría definirse como la congruencia entre las expectativas del contexto y la oferta institucional (dimensión externa) y la congruencia entre la plataforma teleológica de la institución y los recursos y procedimientos que se arbitren para la consecución de aquellos (Villarroel, 1999).

La eficiencia se puede definir como una optimización de los recursos en el logro de las metas.

La eficacia se refiere a la consecución de los objetivos y metas. Sin embargo como señala Villarroel (1999) se suele diferenciar entre efectividad y eficacia. La primera abarcaría el logro de objetivos y metas en el ámbito institucional. La eficacia aludiría más bien a la repercusión que logra la institución con sus productos.

Si así se considera la calidad universitaria, la evaluación institucional estaría dirigida a evaluar la pertinencia, eficiencia y eficacia de dicha institución.

La mayoría de los autores entienden que la evaluación de la calidad de una universidad sólo es posible hacerla globalmente, atendiendo a dichos conceptos y recogiendo tanto los inputs, los outputs y sobre todo los procesos que en ella se desarrollan.

El proceso para dicha evaluación, tal y como se está llevando a cabo en Europa, se articula en dos grandes fases: autoevaluación y evaluación externa. De esta manera el proceso es regulado por la propia institución y a su vez la revisión interna adquiere credibilidad cuando es contrastada por agentes externos.

En este sentido, la evaluación institucional permite dar respuesta, en principio a dos tipos de requerimientos: por una parte, es un instrumento básico para la mejora interna de los niveles de calidad de la universidad; por otro lado, facilita la rendición de cuentas y el análisis de la adecuación de los servicios que prestan las universidades a las necesidades y demandas de la sociedad (Jofre y Vilalta, 1998).

4. EVALUACIÓN INSTITUCIONAL EN EL CONTEXTO EUROPEO

En el Memorándum de la Comisión sobre la Enseñanza Superior, se pone de manifiesto que la mera evaluación de la investigación y del personal docente no basta para garantizar la calidad de la enseñanza y que reviste idéntica importancia la calidad de los programas de estudios, los métodos docentes, los métodos de gestión..., poniéndose de manifiesto la necesidad de inculcar la cultura de calidad, "de las cosas bien hechas" en las instituciones con vistas a los intercambios y colaboración.

Así pues en noviembre de 1991 el Consejo y los Ministros de Educación reunidos en el Consejo se propusieron desarrollar una acción comunitaria en el campo de la evaluación de la calidad de la enseñanza universitaria.

Se consideró, que sin perder la autonomía de los estados miembros, éstos se podrían beneficiar a nivel nacional de las experiencias de evaluación a nivel europeo. De hecho la comisión, presentó un informe de los métodos tradicionales de evaluación de las universidades y la propuesta de nuevos métodos. En principio se establecieron dos proyectos europeos -uno en materia de Ingeniería y otro en materia de comunicación/información o Arte/Diseño- en los que participaron cuarenta y dos centros de los Estados miembros.

El objetivo de los proyectos era aumentar la toma de conciencia sobre la necesidad de evaluar la Enseñanza superior en Europa, enriquecer los procedimientos nacionales de evaluación existentes, añadir una dimensión europea a la evaluación de la calidad, apoyar la transferencia de experiencias y contribuir a la mejora del reconocimiento mutuo de títulos y de periodos de estudios mediante el fomento de la cooperación entre Centros (Buendía, 1999).

Tras dos conferencias de gran éxito sobre el tema de la calidad y la evaluación, el Comité Permanente de la CRE (Conferencia Permanente de Rectores de las

Universidades Europeas), decidió en 1993 ofrecer a sus 500 universidades asociadas la posibilidad de ser auditadas para clarificar y evaluar sus aspectos positivos y negativos en el área de gestión de la calidad.

La CRE pretendía ofrecer un diagnóstico externo desde el punto de vista de dirigentes universitarios con gran experiencia, procedentes de diversos sistemas de enseñanza superior en Europa. Este diagnóstico pondría al descubierto los puntos de calidad y los principales protagonistas en el proceso de toma decisiones de la universidad. La intención del equipo de auditoría no era proporcionar modelos para el desarrollo de las universidades, sino más bien un proceso de auditoría consultiva, es decir una revisión externa de apoyo.

El objetivo a nivel europeo era difundir ejemplos de práctica adecuada, convalidar conceptos comunes de pensamiento estratégico, y elaborar puntos compartidos de calidad que puedan ayudar a las universidades asociadas a reorientar su desarrollo estratégico y que a la vez fortalezca la cultura de calidad en Europa. En este proceso se ayuda a la universidad a:

— Examinar como define sus objetivos a corto y largo plazo.
— Observar los marcos externo e interno que dan forma a su desarrollo.
— Discutir estrategias que incremente su calidad, entendiendo calidad por adecuación de los medios al propósito, sin perder de vista los fines.

Durante los años 1994 y 1995 las Universidades de Göteborg (Suecia), Oporto (Portugal) y Utrech (Países Bajos) encargaron a la CRE que desarrollara la metodología para el programa de Auditoría de la calidad y que lo sometiera a prueba en sus universidades (Barblan, 1995). Este estudio piloto se terminó en enero de 1995. En 1995/1996 tiene lugar una segunda ronda en la que participan diez universidades europeas.

La metodología de trabajo es mixta a través de procesos de autoevaluación y evaluación externa. Se aplicó una Guía de Evaluación, adaptada por cada país. Se presentó el primer informe el día seis de mayo de 1996 en el Consejo de educación.

La experiencia de los proyectos piloto han mostrado que además de conocerse los resultados alcanzados por los centros de educación superior, las medidas de evaluación de la calidad pueden fomentar la capacidad de autorreflexión de los centros y su capacidad para reorganizar sus resultados de modo que pueda satisfacer mejor las necesidades de su entorno económico, social y cultural (Buendía, 1999).

Esto dio lugar a la propuesta de recomendaciones del Consejo, presentado por la Comisión de las comunidades Europeas, relativa a la cooperación europea en materia de garantía de la calidad en la enseñanza superior.

5. EVALUACIÓN INSTITUCIONAL EN ESPAÑA

La preocupación en España por los temas de evaluación institucional comienza hacia 1991, cuando se elevó a consideración una propuesta de recomendación del Consejo relativa a la cooperación europea en materia de garantía de calidad en la Enseñanza Superior.

Anteriormente, en España, la L.R.U. sólo hace mención a la evaluación del profesorado (Art. 45.3), sin tener en cuenta, que es el sistema en su conjunto, el que debe ser objeto de evaluación y mejora.

Ante la constatación de la escasa funcionalidad de las prácticas evaluativas, comienza a surgir un cierto movimiento de reflexión sobre las líneas futuras de acción (Rodríguez, 1998).

Destacamos las jornadas sobre evaluación institucional universitaria, celebradas en Almagro, en las que se reconoció unánimemente la necesidad de la evaluación universitaria para la mejora de la calidad. Fue a partir de este consenso, cuando el Consejo de Universidades constituyó un grupo de trabajo en el que participaron 10 universidades, para comenzar el estudio de un proyecto de modelo de evaluación.

Como consecuencia de este proceso se presentó en el Pleno del Consejo de Universidades, la elaboración del documento sobre evaluación del sistema Universitario Español y se aprobó iniciar un programa de carácter experimental y voluntario de evaluación institucional de las universidades españolas.

5.1. Programa Experimental de Evaluación de Calidad del Sistema Universitario

En 1992, el Pleno del Consejo de Universidades aprobaba una resolución por la que ponía en marcha un Programa Experimental de Evaluación de la Calidad del Sistema Universitario, con el que se pretendía poner a prueba una metodología de evaluación institucional, inspirada en las experiencias internacionales y adaptada a la realidad de las universidades españolas.

Se perseguía evaluar la globalidad del sistema universitario, para lo que se identificaron tres áreas de evaluación: enseñanza (la unidad a evaluar elegida fue la titulación), investigación (la unidad elegida a evaluar fue el área de conocimiento) y organización y gestión (la unidad a evaluar fue la universidad globalmente).

Este programa experimental se desarrolló desde 1992 hasta 1994 y contó con la participación de diecisiete universidades agrupadas en tres niveles:

Nivel I: se pretendía una evaluación global de la universidad. Para ello se recababa una batería de datos estadísticos con el fin de diseñar un sistema de evaluación global. En este nivel de evaluación participaron diecisiete universidades.

Nivel II: con la intervención de seis de las diecisiete universidades participantes, se evaluaron dieciocho titulaciones y treinta y un áreas de conocimiento.

Nivel III: sólo participaron de las seis del nivel II, tres universidades. Este tercer nivel pretendía completar la autoevaluación con la visita in situ de expertos externos y la emisión del correspondiente informe, aunque finalmente no se realizó

dicha visita, ni el informe. Se centró sobre todo en el funcionamiento general de la titulación dentro del contexto global de la universidad.

El proceso que se siguió fue el siguiente:

- Se constituyó un Comité Técnico, formado por miembros del Consejo y expertos en evaluación.
- Se constituyó un Comité de evaluación en cada universidad participante.
- El Comité de evaluación, responsable interno del proceso de evaluación, realizó un autoestudio y elaboró un informe sobre cada unidad evaluada en su universidad.
- El autoestudio fue contrastado por un Comité externo, que visitó la unidad evaluada entrevistándose con las partes implicadas.

Los resultados del programa, presentados en el Informe del Comité Técnico de la Secretaría General del Consejo de Universidades, confirman la adecuación del modelo adoptado, siendo satisfactorio el cumplimiento de los objetivos (80%), a juicio de los expertos internacionales que estudiaron las conclusiones del mismo.

5.2. Proyecto Piloto Europeo de Evaluación de la Calidad de la Enseñanza Superior

Durante el bienio 94-95, se desarrolló el Proyecto Piloto Europeo de Evaluación de la Calidad de la Enseñanza Superior en el que participaron universidades y titulaciones de todos los Estados con la intención de ensayar una metodología general común.

La metodología que se siguió fue idéntica al Programa Experimental, pero el contenido de evaluación se limitó a la enseñanza. Con respecto a España, participaron la Politécnica de Catalunya y Universidad del País Vasco (Ingeniería Superior de Telecomunicación) y las Universidades de Granada y Carlos III (Diplomatura Biblioteconomía y Documentación).

Los objetivos generales que se establecieron en este proyecto fueron (Donaldson, Staropoli, Ottenwaelter, Thune y Vroeijenstijn, 1994, citados en Michavila y Calvo, 1998):

- Sensibilizar a la comunidad académica, y a la sociedad en general, ante la necesidad de los procesos de evaluación en la enseñanza superior.
- Dar una dimensión europea a la evaluación de la calidad, incorporando elementos comunes en los diferentes sistemas de evaluación ya existentes en el seno de la Unión.
- Mejorar y homogeneizar los procedimientos nacionales en funcionamiento en aquel momento.
- Contribuir a profundizar en el reconocimiento entre Estados de los períodos de estudio realizados en el exterior, así como los títulos, impulsando los procesos de movilidad.

Como objetivos específicos se establecieron los siguientes:

- Concentrar el esfuerzo en los métodos de evaluación de la calidad de la enseñanza, aunque se evaluaban también actividades de investigación y gestión, en el grado en que su conocimiento fuese necesario para evaluar la calidad docente.
- Contribuir al intercambio de experiencias entre los distintos países, en el seno de la Unión Europea.
- Permitir la participación de todos los Estados miembros interesados.
- Incorporar en el proceso a la vez programas universitarios y programas de enseñanza superior de tipo profesional.

La mayoría de los expertos (Buendía, 1999; Michavila y Calvo, 1998; Rodríguez, 1998, etc.) coinciden en valorar positivamente la respuesta de participación y el cumplimiento de las fases del proceso.

5.3. El Plan Nacional de Evaluación de la Calidad de las Universidades

En 1995, el Pleno del Consejo de Universidades acordó proponer al Gobierno un Plan Nacional de Evaluación de Calidad de la Universidad. El Real Decreto 1947/1995 de 1 de Diciembre (B.O.E. 9-12-95) aprobaba dicha propuesta.

Se trata de una iniciativa con una duración inicial de cinco años, concretada en forma de convocatorias anuales de proyectos de evaluación.

Los objetivos del Plan son los siguientes:

1. Promover la evaluación institucional de la calidad de las universidades.
2. Elaborar metodologías homogéneas para la evaluación de la calidad integrada en la práctica vigente en la Unión Europea.
3. Proporcionar información objetiva que pueda servir de base para la adopción de decisiones de las distintas organizaciones en el ámbito de su respectiva competencia.

Las características más destacables del Plan son:

- Su clara orientación a la mejora.
- La participación voluntaria de las universidades.
- El objeto de la evaluación no sólo es la enseñanza sino que se extiende a la investigación y a los servicios y gestión de la propia universidad.
- La concesión de ayudas económicas para llevar a cabo el proceso de evaluación acorde con los proyectos presentados por las propias universidades.

5.3.1. *Fases del Plan*

El proceso de Evaluación de la Calidad de la Universidad se articula en tres fases:

1. Autoevaluación.
2. Evaluación externa.
3. Informes Públicos. Comprende los siguientes:
   — Informe de Evaluación de la titulación o unidad evaluada.
   — Informe de la Calidad de la Universidad.
   — Informe sobre los resultados de la convocatoria del Plan Nacional de Evaluación de la Calidad de las Universidades.

A continuación describimos las distintas fases, según la guía de evaluación del Plan Nacional de Evaluación de la Calidad de las Universidades.

A) ***Autoevaluación.*** La autoevaluación es un mecanismo a través del cual, mediante un proceso de reflexión participativa, la comunidad universitaria describe y valora su realidad.

La autoevaluación o evaluación interna, la realiza el Comité de evaluación de cada universidad, de acuerdo con la guía y el protocolo generados por el Comité Técnico.

El proceso de autoevaluación empieza en la recogida y sistematización de información sobre la unidad evaluada. Esta información puede contener estadísticas y datos de gestión, documentos, opiniones y valoraciones originadas en el proceso de evaluación. El análisis se realizará sobre los inputs, los procesos y los resultados, y se plasmará en el autoinforme.

El autoinforme consta de dos partes: por un lado, un análisis objetivo y riguroso de la realidad de la unidad evaluada donde se resaltan los aspectos fuertes y débiles de la misma. Por otro, la propuesta de mejoras en las que se definan prioridad, los procedimientos para poder llevarlas a cabo, y los niveles institucionales implicados en cada una de ellas.

B) ***Evaluación Externa.*** El autoinforme se contrasta mediante una evaluación externa, que realiza un Comité de Expertos Externos, ajenos a la universidad evaluada. Este comité de expertos, después de analizar el Informe de Autoevaluación, visitará durante tres días la universidad, se entrevistará con distintos colectivos y elaborará el Informe de Evaluación Externa. Este informe se remitirá al Comité de Autoevaluación que, una vez analizado, incorporará las recomendaciones y sugerencias al Informe de la Titulación, o unidad evaluada.

C) ***Informes Públicos.*** La síntesis del Autoinforme y del Informe de Evaluación Externa dará lugar al Informe de Evaluación de la Titulación o unidad evaluada que ha de someterse al conocimiento y aprobación del Pleno del Consejo de Universidades.

Al final del proceso, cada universidad reunirá los resultados de las distintas evaluaciones en un Informe sobre la Calidad de la Universidad, donde se expondrán los puntos fuertes y débiles y se indicarán las propuestas de mejora.

El Consejo de Universidades redactará el Informe sobre los resultados de cada convocatoria, dando a conocer públicamente la calidad de las universidades evaluadas.

5.3.2. *Ámbitos de evaluación*

De acuerdo con el Plan Nacional de Evaluación de la Calidad de las Universidades, se propone la evaluación conjunta de las dimensiones de enseñanza, investigación y gestión, con la Titulación como referente en cuanto a las diferentes unidades que serán objeto de evaluación.

A) ***Enseñanza.*** La titulación constituye la unidad significativa en la que pueden analizarse el conjunto de *inputs*, procesos y actividades conducentes a la obtención de un graduado/a universitario/a. El diseño del programa, la organización de la enseñanza, su desarrollo y los resultados obtenidos constituyen el núcleo central del contenido a evaluar.

B) ***Investigación.*** En el ámbito de la investigación se considera como unidad de evaluación el Departamento. Cuando un Departamento contenga varias áreas de conocimiento, se incluirán todas ellas en el autoinforme. La selección de Departamentos a evaluar se hará a partir de las titulaciones evaluadas.

C) ***Servicios.*** En el ámbito de los servicios, las unidades de evaluación que deben considerarse depende de la modalidad de proyecto que se esté desarrollando.

En los proyectos temáticos se evaluarán los servicios directamente asociados a la titulación evaluada, de forma operativa debe hacerse referencia a los servicios administrativos que dependen del Centro de enseñanza donde se imparte la titulación.

En los proyectos globales junto a los servicios que dependen de los centros, se evaluarán independientemente, los Servicios y los procesos de toma de decisión del conjunto de la universidad.

5.3.3. *Organización del Plan Nacional*

El proceso a seguir es el siguiente:

- El Pleno del Consejo de Universidades propondrá al Gobierno las condiciones, requisitos y modalidades de cada convocatoria anual de proyectos de evaluación.
- El Pleno del Consejo designa a los miembros de la Comisión Ejecutiva.
- La responsabilidad directa de la gestión y coordinación de todo el proceso de evaluación recae sobre el Secretario General del Consejo de Universidades. Éste habrá de apoyarse en el Comité Técnico y en la Oficina de Gestión. Son funciones del Comité Técnico elaborar las guías y protocolos que habrán de seguir los Comités de Evaluación (los de autoevaluación y los de evaluación externa) y ofrecer formación a los evaluadores y demás partícipes en el Plan Nacional de Evaluación (González, 1999).

A) ***Comités internos.*** En cada universidad se constituirá un Comité de Evaluación de la Universidad, presidido por el rector y compuesto por responsables y

expertos de la universidad, así como por el director de la Unidad Técnica. Este comité es el responsable de la evaluación y de la redacción del Informe de la Calidad de la Universidad.

El Comité de Autoevaluación lo nombrará el Comité de Evaluación de la Universidad y estará compuesto por responsables de la titulación y personal del mundo académico, administrativo y estudiantil. Este Comité de Autoevaluación es el responsable de elaborar el autoinforme de la titulación que engloba los resultados de la evaluación de la enseñanza, la investigación y la gestión de la docencia y personal implicado en la titulación.

B) ***Comité de Evaluación externa.*** El Comité Técnico del Plan Nacional de Evaluación nombra a un conjunto de expertos para que realicen la evaluación externa. Este comité estará formado por un experto académico, un representante del mundo empresarial y por un experto en metodología de la enseñanza y la evaluación. La figura nº 2 sintetiza cómo se desarrolla el proceso.

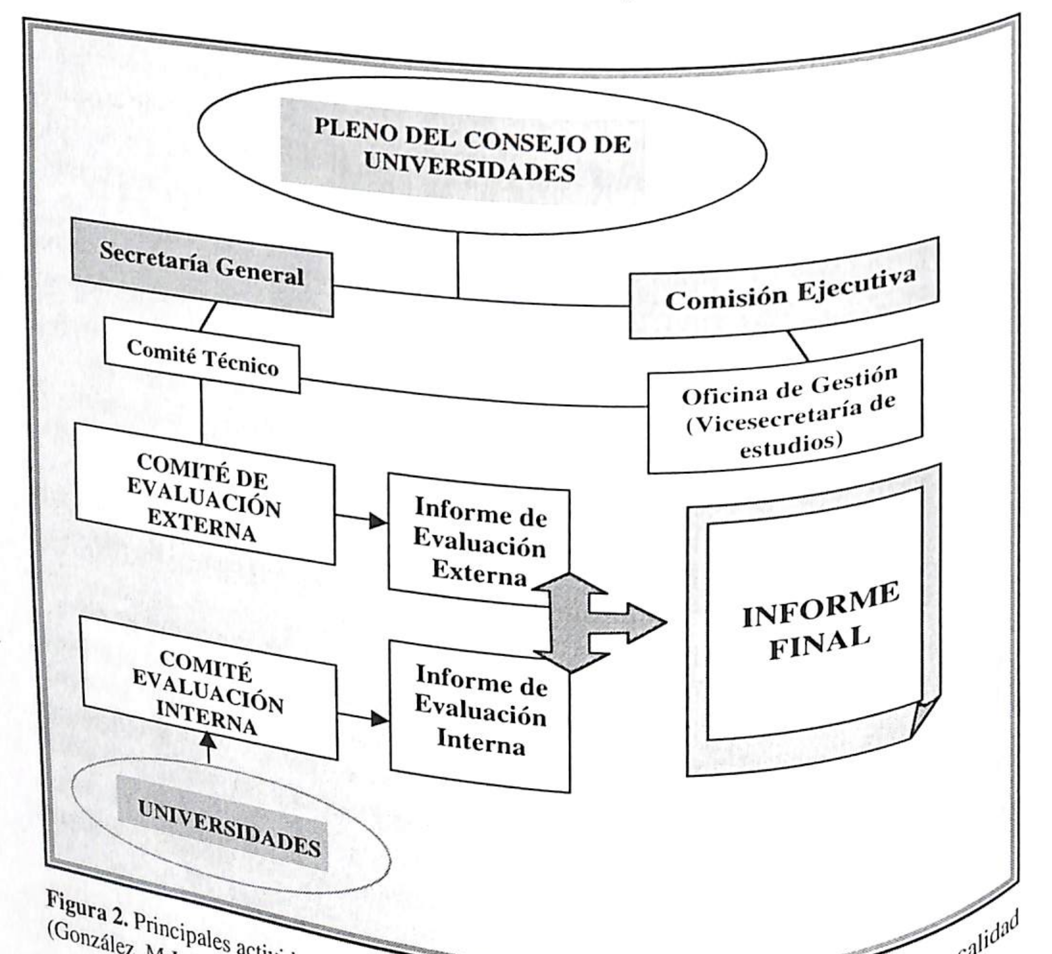


Figura 2. Principales actividades de la educación superior. Plan nacional de evaluación de la calidad (González, M.J. 1999: 110).

5.3.4. *Conclusiones finales sobre el Plan Nacional de Evaluación de la Calidad de las Universidades*

Algunas de las conclusiones a las que han llegado Representantes de las Unidades Técnicas de Evaluación de 42 universidades y de la Agència per a la Qualitat del Sistema Universitaria Catalunya y la Unidad para la Calidad de las universidades Andaluzas (UCUA), reunidos en Almagro los días 18 y 19 de Noviembre de 1999 en el II Foro de Reflexión sobre el Plan Nacional de Evaluación de la Calidad de las Universidades, son las siguientes:

Algunas deficiencias de formación y de apoyo técnico a los Comités de Evaluación de las Titulaciones y a los de Evaluación Externa de las mismas.

Un mayor compromiso en el cumplimiento de los plazos de remisión de los informes finales de evaluación, agilizará más aún la información que nuestras universidades ofrecen a las Administraciones Educativas y a la sociedad.

El contraste entre autoinforme e informe externo en el informe final de evaluación y el desarrollo consensuado de un panel de indicadores, son aspectos metodológicos mejorables del proceso evaluador.

Las dificultades de ajuste de la evaluación de unidades como son los Departamentos recomiendan adaptaciones específicas de la Guía para atender con mayor eficiencia el análisis de la estructura de los mismos dentro de la universidad.

La difusión del material complementario y de las adaptaciones metodológicas creadas por las Unidades Técnicas de las distintas universidades, desde el Consejo de Universidades y las Agencias, posibilitará el rápido conocimiento de estas experiencias al servicio de la mejora de la calidad de los procesos evaluadores.

Necesidad de un plan que aglutine estratégicamente los esfuerzos y de un seguimiento que incluya sistemas de re-evaluación que valoren la eficacia de las políticas de mejora.

Quizás podrían mejorarse las nuevas líneas de actuación, planificación e incentivación de acciones de mejora, con actividades de formación específica sobre estrategias y experiencias en la gestión de la calidad universitaria, dirigidas con atención especial a los equipos directivos y rectorales.

En futuras convocatorias, los procesos de evaluación externa pueden mejorar sensiblemente si se actúa sobre la selección y formación de los Comités de evaluación externos.

En definitiva, la valoración que hacen sobre las convocatorias desarrolladas del Plan nacional es muy positiva, aunque consideran el proceso evaluador necesitado, para su continuidad, de dos aspectos básicos:

El mantenimiento e intensificación del compromiso de los Equipos de Gobierno de las Universidades y de las Administraciones Educativas.

El desarrollo de incentivos para estimular el cambio hacia la mejora de la Calidad.

6. MODELOS DE EVALUACIÓN APLICADOS A LA INSTITUCIÓN UNIVERSITARIA

En los últimos años, la preocupación por la calidad en educación superior ha llevado a los países occidentales a generar toda una serie de sistemas de control de la calidad de las instituciones universitarias, conocidas bajo la denominación de evaluaciones institucionales (Mora, 1991).

Podemos distinguir dos enfoques básicos para las evaluaciones institucionales: uno centrado en aspectos de eficiencia, factores económicos y de control, que hace referencia a la optimización de recursos y otro centrado en aspectos organizacionales del proceso de enseñanza, que hace referencia a la verificación del grado de consecución de los objetivos y que se orienta hacia la mejora de la calidad.

La adopción de un enfoque u otro, dará lugar a los diferentes modelos de evaluación institucional. El modelo de evaluación que se adopte ha de estar enraizado con la realidad de cada sistema universitario.

6.1. Modelo de acreditación

La acreditación implica la evaluación de una institución o programa educativo, por parte de un grupo de expertos externos. Para ello, la institución o programa educativo, solicita su evaluación aportando un autoestudio al grupo de expertos externos que lo evalúan y emiten un juicio público a partir de unos estándares preestablecidos.

El autoestudio como pieza clave del proceso de acreditación, es un método para estimular la autorreflexión cuyo proceso implica una revisión completa de medios, funcionamiento y resultados de la institución con el objetivo de mejorar la calidad de sus programas y servicios (Mora, 1991).

Este autoestudio para que tenga éxito, debe reunir las siguientes características (Kells, 1988):

- — El proceso debe ser motivado internamente.
- — Los dirigentes de las instituciones deben estar comprometidos en el proceso y manifestarlo a todos los componentes de la institución.
- — El diseño del autoestudio debe ser apropiado a las circunstancias de cada institución.
- — El proceso debe contener un verdadero propósito de clarificar y valorar las metas de la organización, con el fin de mejorar la situación.
- — El proceso debe estar bien dirigido por un grupo eficaz.
- — La capacidad de una institución para funcionar efectivamente debe ser estudiada y estimulada.

- — Los resultados del proceso deben presentarse en un informe que sea legible para audiencias de diversos niveles e intereses.
- — El principal resultado del proceso debe ser la puesta en marcha de un sistema de investigación institucional, de autoanálisis y de automejora.

Existen dos tipos de acreditación, uno que acredita la universidad globalmente y otro que acredita programas de tipo profesional específico.

Los procesos de acreditación tienden a ser un proceso burocrático de interés decreciente debido a que las instituciones han interiorizado los mecanismos de calidad, incorporándolos a la práctica habitual. Por ello, en EEUU se ha propuesto últimamente, convertir los procesos de acreditación, en procesos de auditoría de la evaluación, es decir, en procesos de meta-evaluación de la calidad (Mora, 1998).

Conviene señalar que la acreditación no se utiliza para establecer un ranking de universidades, que permite clasificarlas por orden de calidad, aunque éste exista en EEUU y goce de gran prestigio. En España, las autoridades académicas se muestran reacias a participar en un ranking de universidades (González, 1999).

6.2. Revisión de programas

La revisión de programas es un enfoque evaluativo que abarca a la totalidad de programas de una institución y pone el énfasis en los procesos, como estrategia de mejora constante de calidad, lo que la distingue del enfoque de acreditación, orientado a la comprobación del cumplimiento de los mínimos estándares preestablecidos. El hecho de que la revisión de programas se haya aplicado a diferentes tipos de programas e instituciones, incluyendo también una orientación de rendición de cuentas, con participación de agentes externos, ha propiciado el uso del término de evaluación institucional para este tipo de evaluación (Rodríguez, 1997).

La revisión de programas, a diferencia de la evaluación, no se confronta con criterios previamente fijados sino que sólo pretende describir y valorar los antecedentes de la unidad o programa, su rendimiento actual, así como sus planes futuros (Holdaway, 1988).

Los dos propósitos básicos que se le atribuyen son: servir como fundamento para la asignación y reasignación de fondos y fundamentalmente, la mejora de los programas.

Aunque no existe un procedimiento universalmente aceptado para la puesta en práctica de la revisión de programas, Holdaway (1988) y Young (1983), señalan algunos de los pasos que son comunes a todas las revisiones:

- Determinar quién está solicitando la revisión y cuáles son sus perspectivas.
- Identificar los propósitos específicos de la revisión.
- Identificar las principales cuestiones sobre las que hay que centrar la atención.

- Limitar el campo de extensión a la revisión.
- Identificar el proceso que va a utilizarse y los periodos de tiempo que van a ser necesarios.
- Diseñar métodos para recoger y analizar datos numéricos, información y opiniones.

Kells (1988a), distingue las siguientes fases para la puesta en marcha de una revisión de programas:

- Realización de un autoestudio por los profesionales del programa que va a ser revisado.
- Revisión externa por expertos del mismo campo científico que el programa revisado.
- Revisión del autoestudio y del informe de los expertos por un grupo de profesionales de la propia institución.
- Como consecuencia de los resultados de los tres pasos anteriores, se llega a un acuerdo con la dirección de la institución.

Algunos de los aspectos a tener en cuenta a la hora de planificar la revisión de programas podrían ser (Tejedor, 1997):

1) Separar la revisión de los elementos materiales y organizacionales de los elementos personales, que requieren una atención especial. Los aspectos a atender serían:
   — Calidad: del profesorado y de los alumnos; del currículum; de los servicios de apoyo; de los servicios administrativos.
   — Necesidad del programa (para la propia universidad y para la sociedad).
   — Demanda actual y previsible de alumnos y de la sociedad por esos titulados.
   — Costes y beneficios.
2) Diferenciar las ópticas individual e institucional.
   — Desde la óptica individual: revisión de resultados (valoración de los aprendizajes de los alumnos/as); la evaluación docente del profesorado o la evaluación de su productividad investigadora.
   — Desde la óptica institucional: metas y objetivos de la universidad, demanda de la titulación, plan docente, indicadores de rendimiento de los alumnos/as...

6.3. Indicadores de rendimiento

La utilización de indicadores de rendimiento de las instituciones universitarias tiene su origen en las evaluaciones externas de carácter gubernamental. Actualmente la utilización de este tipo de indicadores ha crecido debido a la preocupación por la eficiencia y gestión de la calidad, la rendición de cuentas, etc.

De una manera general se llaman indicadores de rendimiento a las medidas objetivas, de carácter cuantitativo, que nos informan de los logros de una institución (Ball y Halwachi, 1987).

Un indicador de rendimiento se diferencia de otras informaciones de síntesis de gestión, estadísticas, etc. en la medida en que éstas no son relevantes para el logro de objetivos institucionales. Un indicador de rendimiento, debe tener relación directa con la funcionalidad y la productividad institucionales y nos debe permitir tomar decisiones para mejorarlas (Escudero, 1999).

Últimamente, está cambiando la definición de indicador de rendimiento, se está pasando de ser exclusivamente una información numérica de tipo estadístico, a algo más completo que incluya juicios subjetivos y opiniones consensuadas (Mora, 1991), aunque sin perder de referencia los objetivos que los determinaron.

El uso de este tipo de indicadores ha sido y sigue siendo fuente de discusión. La crítica más común hace referencia a la posibilidad de que se puedan construir tales indicadores. Estas dificultades han llevado a establecer indicadores de rendimiento parciales, pero el optimizar las partes no significa necesariamente optimizar el sistema como un todo (Sizer, 1979a).

Pese a los problemas anteriormente señalados, Sizer (1982) defiende el uso de estos indicadores ya que ayudan a detectar con rapidez las tendencias de los cambios en las necesidades de la sociedad, en las preferencias por determinados tipos de estudios, y en la reasignación de recursos.

6.4. Evaluación externa basada en los juicios de expertos (peer review)

Conocida también como revisión por colegas, este tipo de evaluación puede ser solicitada por una universidad para el análisis de una o de todas sus unidades académicas y de servicios, o bien puede ser impuesta por el propio sistema universitario (Rodríguez, 1997).

La evaluación externa es realizada por un Comité visitante, formado por expertos en el área, que al mismo tiempo son colegas. Este Comité visitante cumple una doble función: por una parte, la mejora, relacionada con el feedback que el comité aporta al profesorado, y por otra, la rendición de cuentas, a través del informe público que emite y que da a conocer a la sociedad sobre la calidad de la institución (Vroeijensti, y Acherman, 1991).

Este tipo de evaluación ha sido criticada debido a que la emisión de un juicio por parte del Comité de Expertos incluye un componente de subjetividad.

Ahora bien, como señala Mora (1991), el sistema de evaluación que haga uso de indicadores y revisiones por colegas, puede ser el más adecuado. La ventaja que ofrece el uso de indicadores para establecer unos criterios de asignación racional de recursos parece incuestionable, mientras que la ventaja esencial que ofrece

PAÍS	ÓRGANO/AGENCIA	OBJETIVOS	ÁMBITO	TEMAS	PROCESO
REINO UNIDO	Consejo para la Calidad de la Enseñanza Superior Consejo de Financiación	Metaevaluación Calidad interna Sistema de administración Responsabilidad Control de calidad	Auditorías Institucionales Evaluación de titulaciones	Procesos y políticas específicas para la garantía interna de calidad Enseñanza Investigación (por separado)	Autoevaluación Evaluación por pares Visita de campo Autoevaluación Evaluación por pares Visita de campo
FRANCIA	Comité Nacional de Evaluación	Control de calidad Responsabilidad Evaluación de los contratos con el Gobierno	Evaluación de titulaciones A veces, Evaluación de titulaciones	Enseñanza Investigación Extensión (combinados)	Institucional Cuestionarios y visita de campo Autoevaluación Evaluación por pares Visita de campo
PAÍSES BAJOS	Asociación de universidades y Asociación de Politécnicos	Mejora de Calidad Responsabilidad Autorregulación	Titulaciones	Enseñanza Investigación (por separado)	Autoevaluación Evaluación por pares Visita de campo
BÉLGICA	VLIR	Mejora de Calidad Responsabilidad Autorregulación	Titulaciones	Enseñanza	Autoevaluación Evaluación por pares Visita de campo
DINAMARCA	Centro de Evaluación Gubernamentales	Mejora de Calidad Responsabilidad Autorregulación	Titulaciones	Enseñanza	Autoevaluación Evaluación por pares Visita de campo
SUECIA	Escritorio de los Cancilleres de las universidades	Mejora de Calidad Responsabilidad Autorregulación	Titulaciones	Enseñanza	Autoevaluación Evaluación por pares Visita de campo
FINLANDIA	Ninguna	Mejora de la calidad Responsabilidad Autorregulación	Titulaciones	Enseñanza	Autoevaluación Evaluación por pares Visita de campo
PORTUGAL	Fundación de las universidades portuguesas	Mejora de la Calidad Responsabilidad (rendición de cuentas) Autorregulación	Materias	Enseñanza	Autoevaluación Evaluación por pares Visita de campo
ESPAÑA	Consejo de universidades	Mejora de Calidad Responsabilidad (rendición de cuentas) Autorregulación	Titulaciones	Enseñanza Investigación Gestión	Autoevaluación Evaluación por pares Visita de campo

Tabla 1. Cuadro resumen de los modelos de evaluación institucional en diferentes países. Fuente: Salvador, L. (1997).

la revisión por colegas es que posibilita el dar los primeros pasos en la instauración de un sistema de evaluación institucional.

Cabría preguntarse, ¿qué países han utilizado uno u otro modelo, a la hora de implantar un proceso de evaluación institucional?

En el caso de las universidades británicas, los modelos de evaluación se basan fundamentalmente en indicadores de rendimiento y en evaluaciones externas que pueden propiciar "rankings".

En el caso de las universidades holandesas, se combinan la autoevaluación y la evaluación por expertos externos (revisión por colegas, o evaluación por pares), como ocurre en España.

En Francia, las universidades solicitan su evaluación y el CNE nombra a unos evaluadores externos.

En los Estados Unidos, tanto la acreditación como la revisión de programas, han sido los modelos mayoritariamente utilizados.

Para tener una visión general de la aplicación de algunos de estos modelos en distintos países, hemos presentado en la página anterior un cuadro resumen, teniendo en cuenta las siguientes características: órgano/agencia, objetivos, ámbito, temas y proceso.

ACTIVIDADES PARA EL ALUMNO

1. Reflexionar en grupo sobre las ventajas e inconvenientes que para vosotros plantea la universidad hoy en día. Puesta en común del grupo-clase.
2. Ante esta situación que habéis descrito ¿qué papel pensáis que debe cumplir la evaluación institucional?
3. ¿Cómo definirías una institución universitaria de calidad? Justifica tu respuesta y coméntala en grupo.
4. ¿Consideras adecuado el proceso de evaluación institucional universitaria, que se está llevando a cabo en nuestra universidad?
5. Analiza en grupo los modelos de evaluación institucional que plantea cada uno, recogiendo ventajas e inconvenientes.
6. Establece cuatro indicadores de rendimiento en tu centro.

BIBLIOGRAFÍA COMENTADA

DE LUXÁN, J.M. (Ed.) (1998). *Política y reforma universitaria*. Barcelona: CEDECS.

En este libro, se recogen un conjunto de trabajos que desde perspectivas diferentes, unas de carácter académico y otras desde el terreno de la burocracia o desde el ámbito de la política, abordan el estudio de las características de la universidad, estudian las cuestiones abiertas tras la reforma universitaria.

DE MIGUEL, M.; MORA, J.G. Y RODRÍGUEZ, S. (1991). *La evaluación de las instituciones universitarias.* Madrid: Consejo de Universidades. Secretaría General.

Este libro es una recopilación de artículos sobre evaluación institucional que representan los distintos puntos de vista sobre la evaluación institucional, sobre las ideas básicas en torno a ella, sobre las dificultades para su puesta en práctica y sobre las experiencias existentes en los países en los que ha sido ya recorrido algún camino en este sentido.

GONZÁLEZ, M.J. (1999). *La universidad del siglo XXI. Libertad, competencia y calidad.* Madrid: Círculo de empresarios.

El presente trabajo, constituye un análisis de la situación de la enseñanza superior en España y de la visión de sus debilidades y fortalezas. Así mismo contiene una propuesta de líneas de reforma que mejoren su funcionamiento para adecuar la universidad a las necesidades de la sociedad española.

MICHAVILA, F. Y CALVO, B. (1998). *La universidad española hoy. Propuestas para una política universitaria.* Madrid: Síntesis.

Este libro recoge el primer análisis global que se ha publicado en los últimos años sobre política universitaria en España. En él se analizan diferentes temas, dedicando el capítulo séptimo a la evaluación de la calidad de las instituciones universitarias.

MORA, J.G. (1991). *Calidad y rendimiento en las instituciones universitarias.* Madrid: Consejo de Universidades.

El presente trabajo es un estudio sobre temas relacionados con el funcionamiento de las instituciones de educación superior. Concretamente se investiga aspectos de las instituciones universitarias tales como calidad eficiencia, efectividad, evaluación institucional, etc., haciendo especial hincapié en los mecanismos utilizados en diversos países para la valoración de tales conceptos.

CAPÍTULO VII

EVALUACIÓN Y CALIDAD DE LOS SISTEMAS EDUCATIVOS

Mercedes García García[1]
Universidad Complutense de Madrid

OBJETIVOS

1. Conocer el concepto, funciones y procedimientos de la evaluación de los sistemas educativos
2. Tener una actitud positiva sobre la necesidad e importancia de la evaluación de los sistemas educativos
3. Ser capaces de plantear las líneas generales y elegir los procedimientos más adecuados para desarrollar un sistema de evaluación
4. Comprender la importancia de los sistemas de indicadores educativos y conocer los proyectos europeos de indicadores
5. Ser consciente de los problemas más frecuentes que surgen con respecto al diseño y uso de la información de la evaluación de los sistemas educativos
6. Conocer las fuentes de información más importantes para actualizar la información sobre la evaluación de los sistemas educativos

CONTENIDOS

[1] Facultad de Educación y Centro de Formación del Profesorado. Dpto. M.I.D.E. Universidad Complutense. Rector Royo Villanueva s/n. 28040 MADRID. Tfno: 913946231. E-mail: mergarci@eucmax.sim.ucm.es

INTRODUCCIÓN

Un Sistema Educativo se concibe como un *conjunto de instituciones, programas y acciones que una sociedad política destina intencional y sistemáticamente a la educación y/o a la instrucción de sus miembros, especialmente de los más jóvenes* (García Garrido, 1996: 554). Luego cuando hablamos de sistema educativo estamos refiriéndonos, directamente, a la política educativa desarrollada por un determinado país y, en consecuencia, un punto de referencia constante en el discurso de la evaluación del sistema educativo. De hecho, la evaluación será parte de la política educativa y un instrumento básico para elaborar y reformular las acciones políticas emprendidas respecto a la educación: *una forma distinta y nueva de hacer política y, más concretamente, de explicar las políticas por medio de los logros alcanzados o, eventualmente, fallidos* (O.E.I., 1994).

Es verdad que la educación se piensa como un instrumento de desarrollo personal individual pero no puede descartarse tampoco su papel en el desarrollo económico y social de un país. Al fin y al cabo, parece lógico pensar que cuanto mayor sea el nivel de formación de la población, mayores serán las posibilidades de mantener una posición competitiva en el mercado mundial puesto que sus ciudadanos estarán mejor preparados y mejor podrán contribuir al desarrollo científico y tecnológico del país. Es decir, parece compartida la creencia de que invertir en educación es invertir en el futuro económico a través de la mejora del capital humano de la nación. Y no sólo parece un instrumento de desarrollo económico, también se asigna a la educación un poder de transformación social al ocuparse de los miembros en período de formación, receptivos a las actitudes y valores transmitidos a través de la educación (de hecho, a través del sistema educativo se intenta que los miembros de la sociedad se adapten, ya sea en su vertiente de asimilación o de participación, esto es, sumisa, activa o críticamente). Pero para que sea así es necesario, al mismo tiempo, que el sistema educativo responda a las necesidades sociales (nacionales e internacionales) o mejor, sea capaz de identificar problemas y anticipar las necesidades que en un futuro se demandarán a los miembros de la sociedad. Por consiguiente, todo sistema político contempla junto a la planificación y gestión del sistema educativo, la incorporación de mecanismos de control sobre el funcionamiento o el producto de dicho sistema. Es aquí donde aparece la necesidad y conveniencia de la evaluación del sistema educativo. El término de *pilotaje* o conducción que utiliza Landsheere (1996) para referirse a la evaluación es muy claro: *etimológicamente, la idea de pilotaje, de guía, es inherente a la noción de educación -palabra que viene de ducere-. En efecto, educar es conducir hacia un estado deseable. El guía no solamente debe conocer el camino sino también velar para no separarse de él con el consiguiente riesgo de perderse* (p. 15).

Como se puede deducir, el desarrollo de un sistema educativo y su correspondiente sistema de evaluación, se establece en interacción con el resto de siste-

mas (político, económico, demográfico, administrativo, sociocultural) de un país y dentro de un entorno más amplio que corresponde a los ámbitos filosófico, científico, geográfico, histórico de un estado y momento determinado, como explica D'Hainaut (1988). Todo cambio sustancial producido en una sociedad supone, en mayor o menor grado, una modificación del sistema educativo. Este suceso se observa fácilmente en la España de los años ochenta: cambio político, nuevos valores, nuevas necesidades sociales, nuevas realidades geográficas, ... derivan hacia una reforma total del sistema educativo vigente para acomodarlo a los nuevos tiempos y necesidades. Pero el caso de España no es el único, en la mayoría de países occidentales se han producido una serie de fenómenos similares que han provocado, durante la década de los años noventa, la generalización de la evaluación del sistema educativo. Estos fenómenos se asocian con cuatro componentes de los sistemas políticos actuales: la apertura democrática, el control del gasto público, la competitividad económica y la descentralización. Y a una serie de cambios acontecidos en los propios sistemas educativos actuales, como son la masificación y la diversidad. Veámoslo brevemente:

- La apertura democrática de los sistemas políticos conlleva la necesidad de transparencia en la gestión, funcionamiento y resultados de sus acciones y, en consecuencia, la necesidad de informar a la sociedad sobre los logros educativos conseguidos y la adecuación de sus acciones.
- El control del gasto público en educación, así como la inversión que en ella se realiza y el uso adecuado de los recursos deben justificarse ante los ciudadanos mediante el logro de los objetivos educativos planificados.
- La competitividad económica entre los países desarrollados se fundamenta, entre otras cosas, en el desarrollo de un sistema educativo que logre ciudadanos bien formados, entendiendo la formación como la asimilación de conocimientos y destrezas adecuadas a los cambios estructurales, científicos y tecnológicos actuales en un ámbito internacional, de forma que se logre posicionar al país entre las primeras potencias mundiales, lo cual supone la comparación periódica con los países del entorno para comprobar la posición que ocupa.
- La descentralización de la educación se apoya, principalmente, en la necesidad de acercar la toma de decisiones (administrativas, curriculares, didácticas) allí donde se desarrollan las acciones educativas, de forma que se logre una atención más adecuada al contexto y se den soluciones más rápidas a los problemas o desajustes que puedan surgir. Así, es un hecho generalizado la transferencia de competencias educativas a los estados o comunidades autónomas de un país, al mismo tiempo que se necesita comprobar la consecución de unos mínimos comunes consensuados para todos los ciudadanos del estado.
- La masificación del sistema educativo a partir de la democratización de la enseñanza. El acceso de todos a la educación ha supuesto un logro desde el

punto de vista social pero también un incremento de la cantidad de estudiantes y recursos y, en consecuencia, una mayor dificultad de manejar con eficacia la gestión, el funcionamiento y la solución de problemas de todos y cada uno de sus centros. La forma de solucionar este problema, en la mayoría de las sociedades actuales, ha sido intentar disminuir el número de unidades atendidas creando instituciones intermedias situadas en contextos con características comunes como son las Comunidades o Estados de cada país. En la medida que los sistemas se han masificado y diversificado, se ha observado una mayor heterogeneidad de objetivos, de formación del profesorado, de recursos, de poblaciones de alumnado, … y se ha puesto en evidencia la diversidad de criterios y niveles de exigencia en las calificaciones otorgadas a los estudiantes y con ello ha aumentado la dificultad de poder comparar los niveles de aprendizaje alcanzados por los alumnos. Todo ello ha supuesto buscar nuevos instrumentos que permitan la comparabilidad de los logros del sistema.

- La cantidad y variedad de conocimientos actuales, su actualización o el frecuente descubrimiento de nueva información o, incluso, la apertura de vías para obtenerla dificultan la posibilidad de transmitir una formación enciclopédica única, relevante y actual, válida para todos los ciudadanos. A esto se une la constatación y aceptación de diversas realidades culturales y un creciente respeto de la determinación de los pueblos y el reconocimiento de su idiosincrasia. Todo ello pone en evidencia la diversificación de realidades sociales, culturales y científicas a las que se debe dar respuesta desde el sistema educativo a la vez que garantizar conocimientos, destrezas, valores, actitudes… relevantes y comunes a todos los ciudadanos de un determinado estado para permitir conjuntamente la igualdad de oportunidades y la movilidad social. Es decir, cada estado deberá consensuar y controlar una formación básica común a la vez que dé la oportunidad de una formación diversa adecuada a las diferencias sociales, culturales e individuales de sus comunidades.
- Estas nuevas realidades han provocado una urgente puesta en escena de mecanismos de control de la gestión, del funcionamiento, de los procesos y de los logros de los sistemas educativos y a la cada vez más frecuente institucionalización de la evaluación.

1. ANTECEDENTES DE LA EVALUACIÓN DEL SISTEMA EDUCATIVO

La nueva realidad sociopolítica ha hecho que en las dos últimas décadas se haya incrementado el uso de la evaluación de los sistemas educativos en sus dimensiones de acreditación, control, rendición de cuentas y toma de decisiones. Pero la evaluación no es un hecho nuevo, todos los estados, una vez diseñado su sistema educativo, han intentado controlarlo de una u otra manera, bien por ser una fuente

de transmisión de valores y formación de sus miembros, bien para garantizar las competencias de sus ciudadanos. Es decir, desde hace años, los estados han creado algún mecanismo encargado de la obtención de información y control del sistema ya sea ligado a la realización de exámenes nacionales, al cuerpo de inspección educativa o a instituciones específicas de evaluación, aunque no siempre se ha realizado de una forma sistemática, técnica y comparable.

En las primeras etapas de desarrollo de un sistema educativo, cuando existe poca masificación o cuando la composición del profesorado y del alumnado es bastante homogénea, es posible evaluar el sistema o sus componentes a partir de las calificaciones de los alumnos realizada por cada profesor, o por los informes del inspector de la zona junto con los datos estadísticos de matrícula, repetición y abandono. Además, es posible o, al menos más fácil, disponer de y aplicar unos mismos criterios de evaluación, lo que permite utilizarlos como indicadores de los niveles de aprendizaje alcanzados y, en consecuencia, del logro de los objetivos del sistema. Entonces, los antecedentes de la evaluación del sistema educativo, podríamos situarlos en la evaluación sistemática y comparable de los ciudadanos escolarizados. En la medida que la heterogeneidad y descentralización han sido un hecho, la evaluación ha incrementado su carácter institucional.

Koretz (1992) señala tres momentos en la *evolución de la evaluación de los sistemas educativos*:

A) *Primer período*, durante los setenta, vinculado al *movimiento de evaluación de competencias mínimas*. Se desarrollan un gran número de instrumentos para medir las competencias mínimas de los estudiantes con una doble función: la de comprobar que se cumplían los mínimos criterios para graduarse y la de promocionar entre los distintos niveles. En cualquier caso, la evaluación se utilizaba como una función de acreditación pero, además, pretendía no sólo medir el rendimiento sino también mejorarlo y, aunque medían sólo a los estudiantes, en ocasiones se tomaban estos resultados como base para valorar a los profesores, escuelas o distritos.

B) *Segundo período*, durante los ochenta, asociado al *movimiento de reformas educativas*. La detección de desajustes producidos por la evolución de la realidad social y el inmovilismo escolar, o simplemente la detección de problemas, debilidades o lagunas importantes del sistema educativo en marcha, hace que diferentes países se embarquen en cambios de mayor o menor amplitud de sus sistemas educativos, ya sea para actualizarlos, ya sea para mejorarlos[2]. En este período se asigna una nueva función a la

[2] En el primer caso podríamos situar la reforma global emprendida por España con el cambio del sistema educativo promovido con el desarrollo de la LOGSE. En el segundo caso, la situación americana recogida en el informe *A Nation a Risk: National Commission on Excellence in Education* (1983) donde se manifiesta la situación crítica en la que se encuentra la educación del país poniendo en evidencia los errores de sus estudiantes en destrezas instrumentales y su bajo nivel con respecto a otros países.

evaluación, centrándose en la rendición de cuentas. En consecuencia, se desarrollan tests estandarizados, se amplía el número de destrezas y conocimientos evaluados y se incrementa la dificultad de las competencias en lugar de evaluar las competencias mínimas, característica de la situación anterior. En algunos países, los resultados de los tests se utilizan, básicamente, para favorecer o financiar a las escuelas que obtienen notas más altas en el estado y, al mismo tiempo, para poner en evidencia las escuelas y distritos que muestran unos resultados inadecuados.

C) Tercer *período*, en los años noventa, vinculado al *movimiento de la calidad educativa* y a la descentralización de la educación pública. Uno de los compromisos más importantes de las políticas educativas de la mayoría de países occidentales en los últimos años ha sido el de conseguir sistemas educativos de calidad. Así, la evaluación del sistema se ha diseñado para facilitar la toma de decisiones políticas a la hora de emprender cualquier cambio del sistema educativo de forma que se controle la calidad de la gestión, funcionamiento, procesos y resultados educativos. De hecho, la mayoría de estados ha creado algún organismo de evaluación encargado de recoger periódicamente información objetiva y relevante de su sistema y diseñar la forma de controlar su calidad. Además organismos interestatales se han ocupado de la evaluación de los sistemas educativos para lograr estándares comunes de calidad que permitan la comparación de los sistemas de diferentes países. Todo ello ha llevado a elaborar sistemas de evaluación centrados en la recogida de información de cada uno de los elementos críticos (alumnos, profesores, centros, procesos educativos, administración educativa) y de sus resultados en momentos significativos del sistema.

1.1. Primeras experiencias de evaluación de los sistemas educativos

El primer proyecto de evaluación sistemática y permanente de un sistema educativo, centrado en los resultados de los alumnos, se llevó a cabo en 1969 por la *National Assessment of Educational Progress* (NAEP) en Estados Unidos. En los años sesenta, las competencias en materia educativa eran asumidas por cada estado americano o más específicamente, por las autoridades municipales; por consiguiente, se carecía de un currículum nacional o de un sistema nacional de exámenes. En 1963, Francis Keppel, entonces Comisario de Educación, reconoce la necesidad de informar sobre qué es lo que habían aprendido los alumnos durante sus años de escolaridad. Entonces, se designó un comité de investigación para la evaluación del rendimiento escolar, presidida por Ralph Tyler, donde se elaboró el diseño y se desarrollaron los instrumentos de evaluación iniciales. En 1969 se creó una comisión educativa interestatal con la finalidad de delimitar los intereses edu-

cativos nacionales y en 1978 se promulgó una ley para evaluar periódicamente la educación a nivel nacional como forma de controlar los resultados producidos por los diferentes sistemas educativos del país y rendir cuenta de sus progresos.

Entre las experiencias de evaluación pioneras no se pueden olvidar los trabajos de la *International Association for the Evaluation of Educational Achievement* (IEA). Una organización independiente basada en la cooperación entre centros de investigación, fundada en 1958 por un grupo de investigadores relevantes en el mundo educativo (A. Anderson, B. Bloom, R. Thorndike, T. Husen, G. Mialaret, G. de Landsheere, N. Postlehwaite…) quienes intentarían explicar los logros de los alumnos partiendo de las diferencias que se dan de forma natural entre los distintos sistemas educativos desarrollados en cada país, considerando al mundo como un laboratorio. Iniciaron una serie de proyectos puntuales centrados en los resultados de materias instrumentales, caracterizados todos ellos por contemplar tres características: a) evaluación de los logros educativos de los estudiantes; b) de carácter internacional y cooperativo para comparar las diferencias de aprendizaje entre diversos países; c) uso de metodología cuantitativa. Con el paso de los años, los proyectos empezaron a tener un carácter periódico para estudiar la evolución que se producía en el tiempo en diversas áreas de conocimiento, como los tres estudios realizados hasta el momento en matemáticas y ciencias. En la actualidad se incorporan otros proyectos relacionados con aprendizajes básicos en las sociedades actuales (educación cívica, por ejemplo) o en etapas no obligatorias (Proyecto de Evaluación de la Educación Infantil) o en la revisión y mejora de la metodología de investigación.

2. RELEVANCIA Y USOS DE LA EVALUACIÓN EN LOS SISTEMAS EDUCATIVOS ACTUALES

La evaluación del sistema educativo es un caso particular de la evaluación de programas, y como tal se concibe como un instrumento de feedback y mejora pero vinculado a la política educativa, con un amplio objeto de evaluación (el sistema en su conjunto) y audiencia (la sociedad). Se han utilizado diferentes términos para referirse a la evaluación del sistema educativo; cada uno de ellos implica ligeras matizaciones. Así, D'Hainaut (1988) habla de *regulación pedagógica*, refiriéndose al *proceso para la mejora del funcionamiento y resultados del sistema educativo*. Por su parte, Landsheere diferencia entre *control del sistema educativo* (1985), cuando habla de la *observación sistemática y regular de la evolución de un sistema educativo a fin de mantenerlo en la línea de los objetivos fijados y de descubrir las eventuales disfunciones*, y *pilotaje* (1996) cuando alude a *la toma de decisión a nivel macroscópico sobre la base de hechos, tan objetivos como sea posible, relativos al estado, al funcionamiento o a los productos del sistema*. Para este autor, la evaluación no sería más que una fase intermedia dentro del pilotaje, aquella referi-

da al análisis y valoración de los datos recogidos para facilitar la toma de decisiones. Para D'Hainaut, por el contrario, la evaluación incluiría tanto la adquisición de información (sobre el funcionamiento y los resultados del sistema) como la comparación (con los criterios de funcionamiento óptimo y con los objetivos preestablecidos, respectivamente), si bien la determinación de acciones correctoras y su puesta en marcha sería parte de la retroacción.

En este capítulo tomaremos el término de evaluación para referirnos tanto al proceso sistemático de recogida y análisis de información periódica como a la emisión de juicios de valor sobre la bondad del sistema educativo evaluado y los logros conseguidos. En este caso, la toma de decisiones sobre las acciones de mejora a emprender tras la evaluación recaería sobre los responsables de la planificación educativa, aunque sería una parte necesaria del proceso de evaluación.

2.1. Funciones de la evaluación del sistema educativo

La evaluación del sistema educativo contempla una serie de funciones conectadas entre sí que traslucen, al mismo tiempo, la justificación del diseño de un sistema de evaluación. Puede ser que en algunos países prevalezca una función sobre otra pero, en general, se tiende a utilizarlas todas conjuntamente. En cualquier caso, la evaluación de los sistemas educativos parece ser uno de los instrumentos fundamentales en la toma de decisiones políticas y ciudadanas de un país, de forma que los cambios introducidos pasen a depender de la adecuación a las necesidades y demandas de la sociedad en lugar de responder a intereses corporativos o políticos puntuales. Podemos resumirlas en las siguientes:

- ***DIAGNÓSTICO*** del sistema educativo, es decir, obtener información significativa en un momento determinado, identificando, por una parte, los niveles de logro de los ciudadanos en diferentes momentos significativos del sistema y en diferentes áreas geográficas o sociales y, por otra parte, los problemas que puedan surgir durante su puesta en marcha y funcionamiento. Sería obtener radiografías parciales del sistema para conocerlo mejor y observar su evolución.
- ***FEEDBACK***. Desde un punto de vista sistémico, la evaluación del sistema educativo sería el mecanismo de información sobre sus resultados, procesos y funcionamiento. Tendría una doble vertiente: a) interna o TOMA DE DECISIONES: los responsables de la política educativa del país obtienen la información necesaria para tomar decisiones sobre las medidas o acciones a emprender (de mejora, innovación o cambio), conducir el sistema educativo vigente reorientándolo hacia el logro de los objetivos, adecuar los fines del sistema educativo a las necesidades y demandas sociales, valorar las medidas adoptadas y controlar la calidad del sistema; b) externa o RENDICIÓN DE CUENTAS: dando información sobre el sistema a los ciudadanos, sería una forma en que los responsables políticos rinden cuentas a la sociedad de los resultados, acciones y programas educativos desarrollados y de los gastos/inversiones realizadas en educación. De esta manera, los usuarios podrían juzgar la competencia de sus políticos, opinar sobre las medidas adoptadas o la pertinencia de las estrategias educativas empleadas o, incluso, elegir el mejor centro. Además, podrían pedir responsabilidades y el cumplimiento de promesas a quienes toman las decisiones.

- ***COMPARACIÓN*** de los logros y características de los diferentes subsistemas de un país o entre los sistemas educativos de diferentes países, utilizando sistemas de indicadores comunes. La comparación permite situar la posición en que se encuentra el sistema objeto de evaluación y así, mediante criterios relativos basados en la distancia entre los indicadores elegidos, obtener información sobre los resultados satisfactorios y los mejorables, intentando equiparar o situar el sistema en las primeras posiciones.
- ***CONTROL o ACREDITACIÓN***. Es necesario supervisar que los principios, gestión y componentes del sistema educativo se están poniendo en práctica y se están desarrollando bajo los estándares de calidad formulados. Así, entre otras acciones, deberá realizarse un control de la adecuación de las líneas curriculares prescritas, de la respuesta coherente a los fines sociales y necesidades personales, de los resultados logrados, de los recursos y acciones empleados, etc. que sea útil no sólo a los responsables políticos (en los respectivos ministerios, sindicatos...) sino también a los responsables educativos de los centros (directores y profesores) y a los usuarios (padres y alumnos) de forma que se asuman las responsabilidades institucionales pertinentes, se adopten las medidas de mejora que a cada uno le correspondan y se hagan recomendaciones de medidas para la mejora de la calidad del sistema educativo. En la actualidad, no sería tanto un control de la normativa, aunque también, sino de los criterios de calidad establecidos.
- ***PROSPECTIVA***. La evaluación sobre el comportamiento y resultados del sistema educativo permite analizar las tendencias dentro del sistema, pudiendo predecir posibles cambios y anticipar las necesidades futuras. De este modo se pueden elaborar planes a medio y largo plazo así como determinar las prioridades de las acciones a emprender por la política educativa.

2.2. Utilidad y riesgos de la evaluación del sistema educativo

Las funciones de la evaluación que acabamos de citar, deseables en una sociedad democrática, deberían responder al consenso previamente establecido sobre

los fines de la educación, los criterios de calidad y los resultados deseables, así como también debería estar libre de manipulaciones e intereses particulares. Sólo en la medida que se trabaje con información relevante, fiable, objetiva y válida, la evaluación se convertirá en un instrumento imprescindible de orientación política, centrada en la supervisión, mejora y logro de la calidad de la educación.

La utilidad, sin embargo, va ligada a la asunción de una cultura positiva de evaluación. Debemos evitar los sesgos en sus procedimientos y usos así como las interpretaciones simples y las generalizaciones infundadas a los que la evaluación es proclive. En la medida que se realice por técnicos en evaluación, conocedores de la potencia y límites de sus procedimientos, los planificadores educativos contarán con el mejor mecanismo para emprender cualquier medida de cambio, parcial o total, de forma pertinente y adecuada (por ejemplo, el tipo de jornada escolar).

Los problemas que repercuten, con mayor frecuencia, en una práctica inadecuada de la evaluación, y una consiguiente inutilidad de las informaciones obtenidas, se centran en la forma de evaluar y en el uso que se hace de la misma. Lo primero tiene que ver con las finalidades y el diseño de la evaluación, lo segundo, con el informe y las interpretaciones de la evaluación.

Para desarrollar una evaluación correcta y útil, podríamos resaltar la necesidad de:

- *Controlar el diseño y la implementación del sistema de evaluación*, teniendo en cuenta la necesidad de: a) consensuar qué objetivos son comunes y relevantes así como los momentos y las dimensiones a evaluar. En cualquier caso habría que definir y justificar qué se quiere, debe y puede evaluar. Se debe ser consciente de la dificultad de la empresa y de la necesidad de restringir la evaluación a determinados objetivos de aprendizaje o materias instrumentales o niveles puesto que no todo es mensurable, tanto desde el punto de vista conceptual como económico. Sólo posibilitando la adecuación a la realidad educativa y el acuerdo de lo prioritario se hará viable la evaluación, se conseguirá información relevante y se facilitará la comparabilidad de los datos. Al mismo tiempo, se debe evitar recoger información, única y exclusivamente, de las áreas de matemáticas y lengua o a objetivos conceptuales. Esta restricción corre el peligro de priorizar la enseñanza de unas materias en detrimento de otras áreas, así como dejar de evaluar objetivos asociados a actitudes, valores o procedimientos relevantes en la formulación de un sistema educativo armónico e integral, ofreciendo sólo información parcial o sesgada de la realidad educativa; b) diseñar un sistema de evaluación técnicamente correcto, que utilice medidas relevantes, objetivas, fiables y válidas, evitando todas aquellas fuentes que puedan invalidar los resultados (referidos a la población, instrumentos, condiciones de aplicación, etc.). Además es necesario ser conscientes de las limitaciones de la propia evaluación, más cuando tiene un componente político como ésta y puede ser fácilmente manipulable; c) elaborar un plan a largo plazo que recoja, sistemática y periódicamente, información puesto que una sola evaluación por sí misma no ofrece información suficiente sobre la evolución del sistema, sino que es una foto fija que describe un momento determinado pero que sólo tiene sentido en la medida que pueda ser comparada a lo largo de los años.

- *Evitar problemas asociados al informe de evaluación y a sus interpretaciones*. Proporcionar información comprensible pero completa es un deber de los responsables políticos y un derecho de los ciudadanos, pero es aquí donde más fácilmente puede darse información incompleta o sesgada y, en consecuencia, inadecuada por manipulada. La presentación incompleta de la información puede llevar a conclusiones erróneas y a toma de decisiones inadecuadas. Es fácil ocultar información o dar más importancia a determinados resultados para justificar determinadas decisiones o acciones.
- Cuando los resultados de la evaluación van ligados a una política de incentivos o de ranking escolar, es decir, cuando se prioriza la función de rendición de cuentas por centros frente a cualquier otra, se corre mayor riesgo de proceder a un mal uso de la evaluación. En estos casos es más fácil que los centros utilicen los instrumentos de evaluación para enseñar los conceptos, destrezas o materias incluidos en los items con el fin de mejorar las puntuaciones del centro. Como es evidente, esta práctica, aunque puede repercutir en un aumento de las puntuaciones de las pruebas, no va asociada a la mejora educativa. Dar más peso a los objetivos, contenidos, materias... de las pruebas de evaluación, supone relegar a un segundo plano otras áreas de conocimiento o ámbitos de desarrollo personal que sean educativamente importantes o, incluso, a penalizar a aquellos centros en los que se desarrolla una educación más armónica.

Por último, si los resultados de la evaluación deben promover decisiones de mejora o cambio, alguien o algún organismo deberá asumir la responsabilidad para emprender las acciones más convenientes. Mientras no haya una cultura positiva de evaluación, es posible que se corra el riesgo de que se diluyan responsabilidades y nadie asuma la parte que le corresponde y, finalmente, sean responsabilizados del fracaso o bajos resultados a las familias o a los alumnos, precisamente lo que menos pueden cambiar la situación educativa.

ACTIVIDAD: Recoger noticias de los periódicos sobre las opiniones emitidas por diferentes políticos cuando se publicaron los resultados de la Evaluación de Secundaria y de Primaria. Observar el uso político de la evaluación en función de los resultados logrados por su comunidad. ¿A quién se responsabiliza? ¿Qué medidas se toman?

3. PERSPECTIVAS ACTUALES EN LA EVALUACIÓN DE LOS SISTEMAS EDUCATIVOS

El enfoque de evaluación dependerá del tipo de modelo de gestión política del sistema educativo y de la función que se priorice en la evaluación. Y, en la actualidad, va íntimamente relacionado con el concepto de calidad educativa prevalente.

3.1. Vinculado a la estructura de gestión y funcionamiento

En cuanto al primer caso Tiana (1993, 1997), Oroval (1995) y Toranzos (1997) identifican dos modelos de gestión, funcionamiento y control del sistema educativo: el centralizado y el descentralizado, aunque habría que reconocerlos como los dos polos de una dimensión puesto que pueden darse otros modelos intermedios, últimamente los más frecuentes. Mientras que en el centralizado, prevalece un sistema de acreditación o control, en el descentralizado, un sistema de diagnóstico y toma de decisiones.

Los sistemas educativos actuales se desarrollaron a partir de modelos de organización centralizados. Este modelo se caracteriza por su homogeneidad curricular y por un alto grado de control del estado en todas las fases del proceso educativo. La evaluación se realiza en cada centro educativo y se delega en los responsables más directos del sistema (profesores, directores e inspectores). Su función principal es la de controlar que lo legalmente establecido por las directrices de la política educativa se cumplen; por ejemplo, supervisan que las programaciones contienen los criterios establecidos, que los materiales y procedimientos didácticos que se utilizan son los recomendados y aprobados y los conocimientos establecidos se han enseñado y logrado. Cada responsable tiene un ámbito de evaluación, así mientras que el profesorado certifica que los alumnos han alcanzado los objetivos legalmente definidos, el director supervisa que los docentes cumplen correctamente las funciones formativas establecidas, y la inspección controla que en cada centro se está llevando a cabo la planificación, utilizando los recursos prescritos y logrando los resultados esperados. Al mismo tiempo, se dispone de un sistema de certificación (exámenes de estado o certificados escolares) de los resultados logrados en cada nivel educativo o al finalizar cada etapa educativa del sistema educativo formal. La evaluación, en definitiva, tiene un carácter de acreditación de lo legislado y de corrección y guía hacia las directrices establecidas. Es el modelo que Landsheere (1996) definiría como *pilotaje administrativo* encargado de que se cumplan las disposiciones reglamentarias y los niveles mínimos establecidos por lo que la unidad de análisis será el centro o el distrito escolar.

El modelo descentralizado contempla variedad de opciones políticas y contextuales, caracterizándose por la flexibilidad, diversidad, autonomía y participa-

ción. En este caso cada administración elabora su propio modelo curricular y la forma de gestionarlo basado en sus necesidades y prioridades. Se entiende que la planificación, financiación y evaluación son procesos interdependientes por lo que se necesita elaborar mecanismos de control integrados en el modelo de administración y gestión. Al tener cada administración el suyo propio, este modelo tiene el peligro de diferenciar al máximo a sus ciudadanos en función del lugar donde fueron educados, dificultando la movilidad social.

Habría un modelo intermedio, un movimiento rearticulador en palabras de Toranzos (1997), donde las instituciones centrales deben, por un lado, transfieren la toma de decisiones (administrativas, curriculares, didácticas) allí donde se desarrollan las acciones educativas, de forma que se logre una atención educativa más adecuada al contexto y se gestionen con mayor rapidez las soluciones a los problemas o desajustes que puedan surgir y, por otro lado, deben asumir la planificación, conducción y control del sistema educativo a nivel nacional, aunque bajo nuevos principios: diversidad y equidad. Así, el gobierno central deberá redefinir sus funciones, ocupándose de: a) Orientar globalmente las políticas educativas del país, generando y consiguiendo acuerdos sobre lo común y relevante en la formación ciudadana así como la forma de entender y controlar la calidad; b) Supervisar el equilibrio de los recursos entre diferentes estados o comunidades; y c) Garantizar que todos sus ciudadanos logran las capacidades básicas, socialmente válidas y relevantes, en cada tramo del sistema educativo. En definitiva, deberá centrarse en comprobar la capacidad de sus sistemas para alcanzar los resultados esperados y compensar las diferencias derivadas de desventajas iniciales. Landsheere (1996) lo llamaría *pilotaje del rendimiento escolar* encargado de comprobar el nivel alcanzado y el grado de homogeneidad del producto producido por el sistema para informar a la administración, a la comunidad educativa y a la sociedad. Otra posibilidad sería el *pilotaje diagnóstico o formativo*, vinculado más a una evaluación criterial de las áreas curriculares puesto que compara los resultados obtenidos con los objetivos establecidos para detectar posibles problemas de aprendizaje y buscar soluciones.

3.2. Vinculado al enfoque de calidad educativa

La evaluación del sistema educativo debe iniciarse con la reflexión detallada del estado y situación del sistema educativo vigente y de la formulación consensuada de los principios o metas a seguir. Aunque es responsabilidad de cada país, lo cierto es que en la actualidad, en los diferentes foros internacionales se determinan las líneas educativas deseables a lograr. Su formulación implica a todos los países participantes y así se ha podido observar como, tras el logro durante el siglo XX del principio de universalización de la educación, es decir, de conseguir que todos los ciudadanos tuvieran acceso a la educación, se ha planteado como nuevo reto para los próximos años el lograr una educación de calidad para todos en condiciones de

equidad. Luego, hoy en día la evaluación va ligada al concepto de calidad en educación. Este concepto, aunque fácilmente aceptado, significa como reconoce Muñoz-Repiso (1996: 55) *cosas diferentes según los enfoques y los grupos de interés, según se hable de los resultados o de los procesos, según para quién y para qué sea esa calidad. Está muy relacionada con los fines que se propongan a la escuela, con las políticas educativas de cada país y, en último extremo, con la idea de hombre y de sociedad y con los valores subyacentes a cada sistema educativo.*

Pérez Juste (1998) formula tres enfoques al hablar de calidad educativa: absoluto, relativo e integral. El enfoque absoluto relacionaría calidad con excelencia, luego un sistema de calidad sería aquel que logre *alcanzar la formación integral de la persona con amplios conocimientos, mente crítica, capacidad para entender la realidad social, consciencia de sus deberes y derechos, actitud de compromiso,* etc. El enfoque relativo relacionaría calidad con superioridad, así se hablaría del sistema del de más calidad tras compararlo con lo anterior o con otros diferentes a partir de criterios previamente establecidos. El enfoque integral asociaría calidad con fines y logros, así como con procesos educativos: *diseñar un sistema en el que la calidad de las metas se articule con unos procesos eficientes que den lugar a resultados elevados, de forma que pueda hablarse de niveles adecuados de eficacia. Sólo en tales casos la eficacia merece la pena y puede ser tomada como criterio supremo de calidad.*

En definitiva, la calidad del sistema educativo podríamos definirla y controlarla desde cuatro dimensiones interdependientes: calidad como relevancia, calidad como eficacia, calidad como eficiencia y calidad como equidad.

- *Relevancia:* asegurar que lo que aprendan los estudiantes responda a las necesidades sociales e individuales, desarrollándose a la vez como ciudadano y como persona integral (física, intelectual, afectiva, social). Se centraría en la evaluación de la coherencia entre necesidades sociales y logros. También se ha venido denominando funcionalidad, valor o utilidad (en este último caso también hay un componente de satisfacción de los usuarios).
- *Eficacia:* lograr que los estudiantes aprendan aquello establecido y consensuado en los planes de estudio en el tiempo establecido (etapas, ciclos o niveles). Se centraría en la evaluación de los resultados logrados por el sistema educativo en función de los fines formulados.
- *Eficiencia:* lograr que los medios, estrategias y recursos utilizados favorezcan experiencias educativas enriquecedoras y que permitan, supuestamente, aumentar el nivel tecnológico y económico del país. Se centraría en la evaluación de la racionalización de costes y la relación entre recursos empleados, procesos utilizados y resultados.
- *Equidad:* lograr que todos los estudiantes, cualquiera que sea su origen y condición (personal, familiar o social), obtengan igualdad de oportunidades, procesos y resultados. Se centraría en la evaluación de la homogeneidad de resultados en todas las zonas geográficas y sociales del país o allí donde se observen diferencias en el acceso o procesos educativos (como sucede con el género) para garantizar la compensación de diferencias y la igualdad de oportunidades. El término de valor añadido que se observa en algunas definiciones de calidad, sería el componente que enlazaría las dimensiones de eficacia y equidad.

Asumiendo que las dimensiones de RELEVANCIA y EQUIDAD deben estar presentes cualquiera que sea el planteamiento a seguir, el *modelo de evaluación* variará según se ponga el énfasis en la eficacia o en la eficiencia, entonces:

Si domina la dimensión de EFICACIA, la educación de calidad se entenderá como aquella que logra que los alumnos aprendan "lo que deben" a su paso por el sistema educativo. Lo que hay que dejar claro es el significado de lo que se debe aprender o se espera que los alumnos aprendan, y determinar si se deben evaluar los mínimos en cada tramo o si, por el contrario, deberían evaluarse máximos deseables. En cualquier caso, la estrategia metodológica se centrará en medir periódicamente los resultados del aprendizaje o rendimiento educativo, en distintos momentos críticos del sistema (primaria, secundaria), en diferentes áreas de conocimiento (lengua, matemáticas, ciencias sociales, etc.), en las diversas vías curriculares, etc. en todo el territorio nacional. Los datos se analizan de forma cuantitativa y en algunos casos se hace un análisis posterior de las pruebas, detectando los errores más frecuentes cometidos en las respuestas, identificando los ítems de mayor dificultad en cada área y, finalmente, elaborando propuestas metodológicas a los profesores, a modo de recomendaciones, para la mejora del rendimiento en cada área de conocimiento.

Si, por el contrario, domina la dimensión de EFICIENCIA, la educación de calidad será aquella que proporciona los medios y recursos necesarios a los agentes educativos (equipos de inspección, directores y docentes) para que los procesos de enseñanza y de aprendizaje se realicen en el tiempo y coste más adecuado. Más que en los resultados del sistema educativo, la evaluación se centrará en evaluar las estrategias de gestión y didácticas realizadas en las instituciones educativas. Se plantea para mejorar los procesos, por lo que será necesario identificar los factores asociados a los resultados logrados, relacionándolos con las características de los alumnos y de su medio familiar, de la escuela (modelo de gestión, recursos, clima, metodología docente, etc.) y del contexto social. En este caso, la identificación y medida de factores modificables del proceso educativo permitirá planificar cambios y desarrollar intervenciones sistemáticas y planificadas de mejora mientras que, por el contrario, la utilización e identificación de factores estructurales relacionados con el sistema social, no modificables en el sistema escolar a corto plazo, permitirá emprender acciones de apoyo o programas especiales de compensación para dichas poblaciones, contextos o situaciones.

4. PROCEDIMIENTOS DE EVALUACIÓN DE LOS SISTEMAS EDUCATIVOS

Para garantizar la legitimidad de las acciones de evaluación, el estado debe promover la viabilidad de la evaluación desde varios frentes:

a) Social: creando una cultura de evaluación dirigida a la mejora de la gestión y el control de la calidad de la educación.

b) Orgánico-institucional: garantizando el espacio, los medios, la responsabilidad y la participación de los gobiernos locales coordinados bajo una institución estatal.

c) Técnica: logrando acuerdos y desarrollando equipos técnicamente competentes en las diferentes fases de la evaluación para garantizar la comparabilidad.

Se podrían diferenciar dos *formas de evaluar los sistemas educativos en función del procedimiento* utilizado: estudios estadísticos e investigación evaluativa. No son excluyentes, de hecho todo país en la actualidad mantiene los dos procedimientos y en algunos se añadirían los informes de evaluación elaborados por comisiones socio-políticas, de los que se presentan varios ejemplos en el punto referido a la evaluación del sistema educativo español.

- *Estudios estadísticos*: periódicamente se recogen datos relacionados con el estado del sistema educativo y son publicados por los Institutos de Estadística correspondientes. Los datos no se analizan ni valoran, simplemente se hace una descripción, en términos absolutos y relativos, de los distintos componentes del sistema analizado (número de matriculados en cada etapa, número de alumnos y de alumnas, ratio profesor-alumno, gasto por alumno, etc.) de un año determinado o, normalmente, presentando la evolución de los últimos años.
- *Investigación evaluativa*: se establecen planes de recogida de información sistemática, de carácter transversal o longitudinal, sobre diferentes elementos del sistema con la finalidad de poder comparar resultados entre diversos ámbitos relevantes (currícula, comunidades, tipos de centro, etc.) o para analizar la evolución y valorar las reacciones de las acciones establecidas. En cualquier caso, la evaluación se realiza para orientar (elaborar y reformar) la política educativa y tomar decisiones para mejorarlo o reformarlo.
- Se realizan dos tipos de estudios, unos centrados en la *evaluación de elementos significativos* del sistema educativo (entre los más frecuentes podemos citar: a) Evaluación de los alumnos: recogida de información, a nivel nacional, de una muestra representativa de un determinado ciclo o etapa; b) Evaluación de los recursos: evaluación de los efectos de determinados elementos del sistema, como puede ser la formación inicial y continua del profesorado o de los directores, o del uso de recursos materiales; c) Evaluación de las innovaciones: estudio de las alternativas propuestas junto a las existentes, para que los cambios dependan de una evaluación objetiva; d) Evaluación de los centros de enseñanza: determinar los factores de eficiencia analizando el tipo de funcionamiento de los centros donde se obtienen mayores niveles de rendimiento). Otros estudios se centran en el *desarrollo de un sistema de indicadores* que facilite la evaluación del sistema educativo, de manera que se obtenga información sobre los gastos, el funcionamiento y los resultados del sistema, a la vez que se compara la diversidad de los propios sistemas (área geográfica, líneas curriculares, etc.).

4.1. Fases de la evaluación del sistema educativo

Se podría resumir en cuatro fases la evaluación de los sistemas educativos: planteamiento, diseño, información y toma de decisiones, aunque habría que matizar en cada fase si la evaluación se refiere al control de reformas educativas o a la conducción del sistema educativo vigente (resumidas en la tabla nº 1). Sin embargo, cualquier evaluación debería ir precedida por una correcta identificación de las líneas o principios de la política educativa, en primer lugar, y del problema o planteamiento del objetivo de la evaluación, en segundo lugar. Además, sería el momento apropiado de formular también el modelo de evaluación más adecuado al problema y política evaluadora (de control o acreditación, de resultados o eficacia, procesual o de eficiencia).

4.2. Algunos elementos claves en el proceso de evaluación

Como en toda evaluación es necesario definir claramente los objetivos del estudio. En este tipo de evaluación se partirá de los principios y fines generales del sistema educativo y sus prioridades y, una vez elegido el contenido de la evaluación, se revisarán tres tipos de fuentes documentales: a) legislación y normativa estatal y autonómica; b) datos estadísticos básicos tomados de los respectivos centros de estadística y del servicio de inspección técnica educativa; c) revisión de evaluaciones y resultados previos.

Por último, se consensuará el objetivo a evaluar entre los que se destaquen como fundamentales pero viables, formulando los objetivos y los elementos a evaluar. Un ejemplo lo encontramos en los estudios de la IEA y del INCE, presentados en el recuadro de la página siguiente:

A) ***Determinar el contenido y momento de la evaluación.*** En un estado centralizado, la administración central determina qué evaluar, quién evaluará, cómo evaluar y cuándo, mientras que en el descentralizado, el logro de consenso sobre el contenido de la evaluación es el paso previo para garantizar que la evaluación se

	CONTROL DE REFORMAS EDUCATIVAS	CONDUCCIÓN DEL SISTEMA EDUCATIVO VIGENTE
Planteamiento	Identificación de los problemas, lagunas o pérdida de relevancia del sistema vigente. Propuesta y análisis de alternativas viables tras la puesta en evidencia de los problemas del sistema vigente.	Determinación de los principios educativos prioritarios y metas del sistema político y educativo en el marco nacional, así como los objetivos relevantes de cada etapa educativa. (En un sistema descentralizado esta fase supone el logro de consenso entre los distintos subestados o comunidades autónomas para asegurar la viabilidad política y técnica).
Diseño	Diseño, puesta en práctica y evaluación de proyectos piloto o experimentales con grupos de control equiparables.	Diseño y puesta en marcha del plan técnico de evaluación para períodos quinquenales centrándose en las áreas de conocimiento objeto de evaluación así como en las variables institucionales y socioculturales relevantes. Desarrollo de indicadores, instrumentos y estrategias para la producción de información significativa y comparable.
Información	A los responsables políticos.	A todos los implicados y usuarios.
Toma de decisiones	Abandono, replanteamiento o generalización de la reforma.	Los directamente implicados (políticos, sindicatos, directores, profesores) asumen sus responsabilidades, priorizando las acciones de mejora y supervisando la puesta en práctica.

Tabla 1. Fases de la evaluación del sistema educativo.

ACTIVIDAD *Reforma:* Lee y resume la evaluación de las enseñanzas medias realizada por el CIDE (Álvaro et al., 1988, 1990, 1992). Teniendo en cuenta los resultados encontrados: ¿era aconsejable reformar? ¿se podían anticipar problemas? ¿el diseño fue técnicamente correcto? ¿eran equiparables las poblaciones comparadas?
ACTIVIDAD *Conducción:* Lee y resume la evaluación de la Educación Primaria realizada por el INCE (1995). Identifica los objetivos de la evaluación, el modelo utilizado y los puntos fuertes y débiles detectados en esta etapa educativa. ¿Qué acciones de mejora podrían sugerirse?

EJEM. A: TIMMS (IEA, 1997)
OBJ. 1: Conocer el nivel de rendimiento de los alumnos.
OBJ. 2: Comparar los resultados entre países.
OBJ. 3: Explicar las diferencias observadas en función de características de los sistemas educativos.
EJEM. B: Diagnóstico Educación Secundaria (INCE, 1998)
Describir la situación de la educación secundaria en España relacionada con:
- Qué aprenden los alumnos.
- Qué se enseña a los alumnos.
- Quién enseña a los alumnos.
- Dónde y en qué condiciones aprenden los alumnos.
- Cómo colabora la sociedad con el sistema educativo.

realizará en condiciones adecuadas. Las áreas de conocimiento en las que es más fácil ponerse de acuerdo sobre la necesidad de evaluación son: lectura, escritura, matemáticas, ciencias naturales, ciencias sociales, lengua, literatura y, en menor medida, educación física, arte, música y desarrollo técnico o profesional. Sin embargo, como la sociedad suele otorgar más importancia a la lectura y a las matemáticas, suelen ser éstas las áreas más frecuentemente evaluadas (anual o bi-anualmente) mientras que el resto se realiza cada 5 ó 6 años.

Algo más difícil, pero necesario, es lograr acuerdos en los objetivos y criterios de evaluación. El procedimiento de decisión puede ser encargado a una comisión permanente, representada por gestores y técnicos de cada gobierno local, a una agencia externa o, incluso, se puede intentar ampliar la responsabilidad de determinar el contenido de la evaluación a otros sectores sociales como sucede en el caso de la NAEP. Este organismo intenta obtener una lista de objetivos en base a dos criterios: a) identificarse como relevantes por expertos y usuarios y b) enseñados en las escuelas. Para conseguirlo utiliza el procedimiento siguiente:

i) Grupos de especialistas de las áreas objeto de evaluación determinan los objetivos educativos que los alumnos deben alcanzar en ese campo.
ii) Grupos de directores escolares y profesores determinan si dichos objetivos han sido enseñados.
iii) Grupos de padres y personas no relacionadas con la educación valoran también esos objetivos (acompañadas de algunas preguntas y actividades para ejemplificar el objetivo) por medio de dos cuestiones: ¿qué es lo que cree que es importante que se enseñe en las escuelas? ¿qué le gustaría que su hijo aprendiese?

Lo más frecuente es que expertos en las áreas curriculares y en medida educativa junto con representantes institucionales, determinen objetivos y contenidos básicos a evaluar. Cualquiera que sea el procedimiento, es necesario, a la vez que se logra el consenso de las metas y objetivos comunes, se debe determinar el momento de dominio en el que la mayoría de estudiantes deberían haber alcanzado los objetivos. Dos alternativas son posibles y en ambas se evita recoger información por curso académico: a) evaluar al final de cada etapa o período educativo significativo en vez de por niveles porque cada escuela puede desarrollar diferentes secuencias de aprendizaje en cada materia, como sucede en los estados cuyos currícula sean abiertos y, en cambio, son prescriptivas las capacidades a lograr al finalizar cada etapa; en estos casos si se recogiera la información por nivel educativo obtendríamos una información sesgada del aprendizaje logrado puesto que una escuela podría emprender el aprendizaje posteriormente a otra; b) determinar la edad en la que se piensa que los estudiantes deberían haber alcanzado determinados conocimientos, destrezas o capacidades (como leer, dividir o comunicarse en otro idioma). En el caso de la NAEP, por ejemplo, se utiliza la edad en lugar del curso porque las escuelas americanas utilizan diferentes criterios de promoción y de exigencia. Así,

para obtener información sobre la educación formal obligatoria se especifican tres momentos de evaluación: a los 9 años (cuando la mayoría ha finalizado el tercer curso y supuestamente han adquirido las destrezas instrumentales básicas), a los 13 años (cuando la mayoría ha terminado la enseñanza elemental) y a los 17 años (final de la enseñanza secundaria). Además se evalúan los resultados de la educación no obligatoria recogiendo información de los jóvenes entre 26 y 35 años.

Como resultado de este proceso se obtiene una base de items representativa para cada objetivo y grupo de edad o área de conocimiento o ámbito a evaluar. Algunos de estos items serán incluidos en las evaluaciones posteriores para analizar la evolución a lo largo de los años.

B) ***Elegir los sujetos de evaluación.*** Obtener puntuaciones de todos los alumnos de un grupo de edad supone un enorme costo, por ello es necesario utilizar procedimientos de muestreo que abaraten el gasto sin perder representatividad y fiabilidad en la información recogida. En cualquier caso siempre podría ser posible plantear el muestreo censal o de toda la población censada en el tramo evaluado. Las ventajas e inconvenientes se resumen en la tabla siguiente:

	VENTAJAS	DESVENTAJAS
CENSAL	• Se obtiene información de todos los individuos • Aumenta la posibilidad de comparación entre aulas equivalentes • Base de datos para investigación pedagógica • Diseño de programas de desarrollo en aulas con desventaja identificadas	• Elevado costo (humano y material): impresión, distribución y devolución, examinadores
MUESTRAL	• Costo menor en los variables (los fijos se mantienen: ítems, instrumentos, manuales, trípticos, equipo técnico y administrativo, aplicadores....) • Se puede informar antes al disminuir el volumen de información procesada	• Los resultados son estimados y su fiabilidad depende del muestreo • Se diluye la información de aulas, centros, ...

Tabla 2. Procedimientos de muestreo.

Sin embargo, normalmente, se utiliza una muestra estratificada en función de las características de la población y de los propósitos de la evaluación[3]. Así, en primer lugar, se debe delimitar el área geográfica y el número de matrícula para seleccionar una muestra probabilística de escuelas representativas de la nación. Un segundo momento consiste en seleccionar al azar a los alumnos que están en el

[3] Una justificación y descripción detallada del procedimiento de muestreo puede consultarse en el segundo tomo del *diagnóstico del sistema educativo: los resultados escolares*, pp. 199-211.

mismo nivel educativo o que tienen la misma edad y están matriculados en esa escuela. Por último, una nueva aleatorización se realiza para que se repartan equilibradamente todas las pruebas comprendidas en una materia, de forma que cada alumno de la muestra responda sólo a una parte del total y, de este modo, no se utilice más tiempo en responder la prueba que lo que se dedicaría a una clase. Este procedimiento permite informar de los resultados comparando:

- Regiones o Comunidades Autónomas.
- Tipos de comunidad (rural, metropolitana y urbana) o tamaño de la comunidad (gran ciudad, suburbios de la gran ciudad, ciudad mediana, población pequeña) o hábitat (menos de 10.000 habitantes; hasta 50.000, hasta 500.000, más de 500.000, grandes poblaciones). .
- Tipos de centro (público o privado).
- Nivel educativo familiar (sin escolaridad superior, escolaridad superior, universitaria).
- Etnias o grupos sociales.
- Sexo de los alumnos: masculino y femenino.

C) ***Establecer los procedimientos de medida.*** Unos estados utilizan pruebas de referencia normativa para permitir comparaciones del rendimiento a nivel nacional. Otros utilizan pruebas de referencia criterial, diseñadas para comparar estudiantes (y sistemas) con niveles absolutos de rendimiento. Otros utilizan ambos. En uno y otro caso supone aplicar los instrumentos a todos los sujetos de la muestra seleccionada.

Pruebas de referencia normativa	*Pruebas de referencia criterial*
Se elaboran pruebas de acuerdo a criterios de fiabilidad y validez para cada nivel de dominio mediante puntuaciones normalizadas que permitan situar a los grupos en una determinada posición con respecto a la población de referencia, se presentan las diferencias entre los sujetos evaluados.	Se definen claramente los objetivos, destrezas y competencias que se suponen representativas, relevantes y válidas para el segmento a evaluar y los respectivos niveles del logro satisfactorio antes de construir el instrumento para evaluar las competencias de los sujetos en diversas áreas y conocer en qué medida los alumnos han alcanzado los niveles esperados.
Los resultados dependen del nivel alcanzado por el grupo. Los resultados se comparan con la población de referencia mediante puntuaciones tipificadas (percentiles, puntuaciones típicas normalizadas, etc.).	Los resultados dependen del nivel absoluto de calidad definido. Los resultados se comparan con los niveles prefijados en cada objetivo en términos del porcentaje de respuestas correctas y el punto de corte establecido para la etapa educativa donde se evalúa.

Tabla 3. Procedimientos de medida.

El formato puede ser el mismo; normalmente se han utilizado items de respuesta múltiple o de respuesta breve por la facilidad y rapidez de corrección. Aun-

que cada vez se utilizan más los items de respuesta abierta, sobre todo para evaluar procedimientos como, por ejemplo, la realización de experimentos, la explicación de fenómenos o la composición escrita. Las pruebas de referencia normativa y criterial tienen la ventaja de poder realizar interpretaciones y comparaciones por centros educativos, siempre que se administren los instrumentos a todos los sujetos. Los resultados dependerán de que los items sean fáciles o difíciles para la población seleccionada. En este caso, sólo tiene sentido interpretar el resultado final y no dar resultados por items individuales. La desventaja es que se limita al número de aspectos a evaluar y que la utilización de la misma forma o de formas paralelas puede llevar a centrar la enseñanza únicamente en aspectos medibles o a la enseñanza de los items de los tests. Para evitarlo se han formulado otras propuestas como: el muestreo matricial, el portfolio o la teoría de respuesta al ítem.

El *muestreo matricial* intenta comprobar el rendimiento de un gran número de estudiantes pero sin necesidad de administrar el mismo instrumento ni, por consiguiente, los mismos items a todos los sujetos. En consecuencia, el banco de items puede dividirse en subconjuntos y administrarse a diferentes estudiantes; por ejemplo, la NAEP ha utilizado este procedimiento dividiendo en tres partes la totalidad del test y posteriormente, aplicando a cada estudiante, aleatoriamente seleccionado, una sola parte. La ventaja, como puede deducirse, es poder ampliar el número de áreas y de objetivos a medir puesto que no tendríamos que limitar el número de items a la cantidad de material que un estudiante individualmente puede responder en el tiempo adjudicado. Al administrarse un mayor número de items, se ampliaría la información recogida e, indirectamente, evitaríamos correr el riesgo de que la enseñanza se centrara exclusivamente en el contenido de los tests. Sin embargo, mediante este procedimiento, no podríamos comparar los resultados entre centros ni entre sujetos, hecho que no es muy importante si la evaluación se realiza para controlar el nivel alcanzado por la población (Koretz, 1992).

La *teoría de respuesta al ítem* (IRT) se centra, de ahí su nombre, en el ítem como unidad de análisis del test en lugar de las puntuaciones totales, lo que libera, en gran medida, que los resultados dependan de las características técnicas del test y de las de la población a la que se administra. La IRT utiliza la dificultad relativa de los items para asignar la puntuación al estudiante, de forma que un estudiante que haya respondido el mismo número de preguntas que otro estudiante pero que uno responda las cuestiones más difíciles y el otro, por el contrario, las más fáciles, tendrían una puntuación diferente. Permite aplicar distintos subconjuntos de items sin perder la capacidad de comparación entre los sujetos porque utiliza items de anclaje que son comunes a todas las formas de las pruebas. No obstante, los resultados no pueden darse a título individual puesto que cada estudiante responde a diferentes formas del test. Los resultados pueden describirse en escalas de rendimiento, permitiendo referencias criteriales en distintos puntos de la escala en función de lo que los sujetos saben o saben hacer. Un ejemplo lo encontramos en la evaluación de los

resultados de la educación secundaria realizado por el INCE (1998) o en algunas evaluaciones del National Center for Educational Statistics (NCES).

Cualquiera de las modalidades anteriores supone crear situaciones artificiales para obtener la información base de la evaluación. El *portfolio*, precisamente, intenta recoger muestras de trabajos escolares realizados en contextos naturales. Su fuente de información es la producción cotidiana que los estudiantes realizan durante períodos largos de tiempo (el curso académico, por ejemplo), recogidos en ficheros o portfolios. Se puede analizar no sólo la producción en un momento dado sino, principalmente, la evolución del aprendizaje durante un período de tiempo. Sin embargo, todavía no se ha llegado a un adecuado nivel de estandarización de las puntuaciones; además, la cantidad de material que se tiene que analizar para cada sujeto, impone restricciones sobre el tipo de tareas a evaluar y el tamaño de la muestra.

D) ***La determinación de los procedimientos de difusión de la información.*** Si uno de los objetivos de la evaluación es ofrecer información a la sociedad y a los responsables de la política educativa, audiencias que no tienen por qué tener formación técnica, entonces se ha de hacer un esfuerzo para que los resultados se interpreten con facilidad, (sin embargo, la sencillez no debe ir en detrimento de la amplitud ni objetividad de la información), por eso los informes no suelen ser excesivamente técnicos. Los resultados suelen presentarse en porcentajes porque existe la falsa creencia de que se entienden mejor que los percentiles, las puntuaciones típicas o las de dominio.

Normalmente se elabora un informe de los resultados de evaluación por cada asignatura, informando del porcentaje medio de respuesta correcta para cada objetivo general, así como del porcentaje de respuesta correcta para cada grupo de objetivos o competencias evaluadas.

Sin embargo, cada vez es más frecuente encontrar los resultados en escalas de rendimiento con puntos de referencia criterial, como el caso del informe que presentamos en la página siguiente, sobre los resultados de la educación secundaria obligatoria realizado por el INCE (1998).

Si las condiciones de procedimiento y uso se cumplen, un sistema de evaluación nacional puede proporcionar una información descriptiva relevante sobre los resultados de sus sistemas educativos, identificando grupos de competencias, escuelas, distritos y comunidades exitosos o aquellos que, por el contrario, tienen dificultades en determinadas áreas o competencias y a partir de esta información, determinar las acciones educativas más pertinentes y establecer las prioridades.

La inferencia sobre la calidad de profesores, escuelas o sistemas eficaces o ineficaces, basándonos en los datos recogidos debe realizarse con prudencia porque: a) la información recogida en un momento determinado nos permite diagnosticar la situación del ámbito y objetivos evaluados, pero no supone el conocimiento de las relaciones causales ni de los factores causantes de los resultados; b) en educación no pueden establecerse todas las fuentes de variación significativas asociadas a los

NIVEL 150:
- Comprende el significado de enunciados específicos de textos utilitarios.
- Integra información del texto para resolver problemas presentados en textos informativos.

NIVEL 200:
- Comprende el significado de palabras específicas, reconoce el tema o idea principal y establece la secuencia temporal de textos utilitarios.

NIVEL 250:
- Comprende el significado de palabras abstractas u otras poco habituales.
- Comprende el significado literal (idea central, personaje, tesis del autor), interpreta las ideas principales, y valora elementos retóricos de textos informativos.
- Integra la información de textos informativos y literarios para obtener una nueva y reconoce el doble sentido.

NIVEL 300:
- Comprende ideas secundarias y enunciados sintácticos complejos de los textos informativos.
- Reconoce e interpreta el sentido figurado y el doble sentido de textos informativos y literarios.

NIVEL 350:
- Comprende ideas abstractas de textos informativos.

NIVEL 400:
- Comprende ideas abstractas en textos utilitarios e informativos.
- Reconoce y valora la precisión de las ideas y las tesis defendidas en textos informativos.
- Interpreta el significado de procedimientos retóricos empleados en textos literarios.

resultados, además de que pueden ir asociados también a factores no escolares o no medidos; y c) la mayoría de los diseños son de carácter transversal, en cambio las influencias educativas y no educativas no suelen producir efectos inmediatos sino que, frecuentemente, son remotas en el tiempo (otros cursos, escuelas...). En definitiva, para mayor objetividad y utilidad se aconseja el conocimiento de los límites de la evaluación y la prudencia en la interpretación de los resultados.

5. ESTABLECIMIENTO DE INDICADORES DE LA CALIDAD DE LA EDUCACIÓN

Los indicadores surgen en el campo económico y social para presentar datos relacionados con ciertas características relevantes de la población o del desarrollo económico del país y, recogiéndolos periódicamente, informar de su evolución (demografía, producto interior bruto, etc.). Posteriormente, los indicadores aparecerán en el campo educativo para informar de los logros y/o deficiencias del sistema educativo, y para observar su evolución. En la actualidad, los indicadores en educación se utilizan de tres maneras:

- Estadísticos descriptivos de recursos y resultados del sistema educativo a nivel macro para conocer qué se pone en el sistema y qué produce.
- Sistemas de indicadores globales sobre contexto, resultados y/o procesos a nivel macro, intentando poner en relación factores de entrada y proceso del modelo con los resultados.

- Sistemas de indicadores sobre la escuela y procesos educativos particulares a nivel meso, intentando identificar escuelas eficaces.

En cualquier caso, la selección y el uso de los indicadores está ligado al momento de desarrollo del sistema educativo de un país. Así, en los primeros estadios de implantación de un sistema educativo, cuando la escuela sólo atiende a una pequeña parte de la población, y se piensa que por el simple hecho de asistir a la escuela los estudiantes aprenden, entonces, suele ser prioritario, políticamente, el aumentar la cantidad (de ciudadanos que asisten, de tiempo de asistencia y de recursos) para que el país obtenga beneficios de la educación, suponiendo unos ciudadanos mejor preparados. En estos casos, indicadores como la evolución de la matrícula, la cobertura de la escolaridad obligatoria, la repetición o el abandono, ofrecen información suficiente sobre el sistema educativo del país. En consecuencia, que cuantos más ciudadanos estén matriculados, más finalicen los estudios secundarios -o mejor universitarios-, y menos abandonen los estudios, más sano se supondrá el sistema educativo de un país.

En la medida que se ha generalizado la enseñanza obligatoria y se ha ampliado a un mayor número de ciudadanos, los indicadores que acabamos de mencionar sobre cantidad de usuarios o recursos, tienen cada vez menor capacidad informativa. Hoy en día se es consciente de que el sistema educativo debe, además de ofrecer un lugar, procurar una educación de calidad. Por ello, en la actualidad, en la mayoría de países industrializados, la calidad educativa (su forma de concretarse) es la que rige la selección y desarrollo de los indicadores educativos para informar de qué, cuánto, cómo y en qué condiciones se aprende en las escuelas.

Los indicadores educativos deben ofrecer información relevante sobre el sistema educativo con el objetivo de orientar la toma de decisiones de los responsables de las políticas educativas (o planificadores escolares) y plantear acciones de mejora. Es decir, con los indicadores educativos se pretende:

- Describir y comparar diferentes sistemas educativos.
- Analizar cambios o la evolución de un sistema educativo, con lo que se intenta:
  - — Plantear hipótesis explicativas del cambio.
  - — Predecir cambios futuros.
  - — Estudiar el impacto de reformas o decisiones políticas.
- Informar sobre los problemas, necesidades o potencialidades del sistema educativo.
- Informar a la sociedad de los resultados, recursos,... obtenidos o empleados en la educación.

5.1. El concepto de indicador

Un indicador puede definirse de dos maneras. Una más restrictiva, entiende que el indicador es un *estadístico (dato cuantitativo) referido al sistema educativo*

que revela algo sobre su funcionamiento o salud (J. Oakes, 1986). Una definición más laxa, define el indicador como cualquier variable que sea clara y consistente y que ayude a comprender mejor el nivel alcanzado y/o los cambios producidos en el sistema (Jaeger, 1978) pudiendo ser tanto un dato cuantitativo como un informe descriptivo o narrativo. Desde este punto de vista, un indicador será aquella variable que proporcione información relevante sobre el sistema educativo en términos de valor añadido o cambio de estatus de cualquier grupo de personas, instituciones, elementos o condiciones analizadas.

Un indicador no describe ni explica la estructura del sistema pero los datos que produce sobre la educación son útiles y necesarios para la toma de decisiones sobre las acciones educativas a emprender y pueden revelar la tendencia y la evolución del sistema (Selden, 1990) tomando decisiones a medio y largo plazo. Shavelson et al. (1991) ofrecen una lista de lo que los indicadores pueden y no pueden hacer que pudiera clarificar el concepto de indicador:

LOS INDICADORES PUEDEN	*LOS INDICADORES NO PUEDEN*
• Ayudar a tomar decisiones sobre la política educativa del país, informando sobre los objetivos. • Describir y establecer los problemas más clara y rápidamente. • Describir el estado de la sociedad y sus dinámicas a partir de indicios objetivos. • Valorar las decisiones políticas sobre las acciones emprendidas.	• Determinar objetivos y prioridades. Es un elemento más pero no el único. • Evaluar programas. No proporcionan el nivel de detalle ni profundización necesario para tomar decisiones sobre un determinado programa. • Desarrollar un balance económico (similar al PIB). No son comparables a los indicadores económicos construido bajo una perspectiva monetaria.

La segunda definición de indicador se apoya en que un sistema educativo tiene condiciones o características no medibles directa o perfectamente a través de datos cuantitativos, además de que muchas de las estadísticas que hay sobre el sistema educativo dan poca información sobre su funcionamiento. De ahí que Shavelson, McDonnell y Oakes (1991) propongan una solución intermedia, definiendo el indicador educativo como una estadística simple o compuesta relacionada con un constructo educativo básico que sea útil en el contexto político. Si bien parece haber consenso en que un indicador es una estadística que refleja un aspecto importante del sistema educativo, no por ello todas las estadísticas son indicadores. Sólo se hablará de indicador cuando informe de algún aspecto relevante del sistema y cuando ponga en evidencia cómo está funcionando, cuál es su salud (por ejemplo, el número de estudiantes matriculados es un dato importante, sobre todo en los tramos no obligatorios, pero ofrece más información el porcentaje de alumnos que han completado con éxito y competencia dicho tramo).

5.2. Tipos de indicadores

Habría que distinguir entre indicadores simples e indicadores compuestos. El **indicador simple** es la cuantificación de un solo aspecto o una medida individual del sistema (porcentaje de aprobados en la selectividad, por ejemplo). Pero en muchos casos, un único dato es insuficiente para informar del funcionamiento de un elemento del sistema, por lo que se hace preciso combinar varios estadísticos en un **indicador compuesto**, como es el caso de un buen profesor o la adecuación de un método de enseñanza. Y éste suele ser el caso más frecuente, porque la educación, o la calidad educativa, es un fenómeno demasiado complejo como para ser resumido en un solo dato.

Con la intención de proporcionar una información más relevante sobre el sistema educativo es importante no formular un listado de indicadores, sino agruparlos en factores que les den sentido y den una visión significativa sobre el estado de la educación a los responsables políticos. En este sentido, junto al manejo de indicadores simples y compuestos, se hablaría de un **sistema de indicadores**, proceso que intenta superar la obtención de una mera suma de datos. *Un sistema de indicadores mide distintos componentes del sistema y, además, proporciona información sobre cómo trabaja cada componente individual para producir juntos el efecto final* (Shavelson, McDonnell y Oakes, 1991). Un sistema de indicadores debería reflejar aspectos relacionados con los resultados educativos en términos de incremento (rendimiento educativo, formación pedagógica del profesorado de secundaria, ampliación de los recursos educativos, etc.), relacionándolos con inputs educativos (características de los estudiantes, características del contexto donde está la escuela, recursos económicos y humanos, etc.) y procesuales (adecuación del currículum e instrucción recibida por los estudiantes, clima escolar, organización implicada en la excelencia, utilización de estrategias de atención a la diversidad, etc.).

Desde esta perspectiva los indicadores dejan de ser una información periódica de los distintos inputs y resultados educativos para integrarse en un sistema global, en el que cada indicador interacciona con los demás, de forma que un cambio en uno de ellos supone modificaciones en los otros. El sistema de indicadores pretende informar de los efectos directos de las políticas educativas sugiriendo, a posteriori, posibles explicaciones de los cambios observados en los resultados. En definitiva, un buen sistema de indicadores deberá proporcionar información fiable y precisa sobre la educación para ayudar a tomar decisiones políticas a la hora de formular las metas educativas, establecer prioridades y emprender acciones de mejora. Un ejemplo es el documento *goals 2000* formulado en la cumbre de 1989, convocada por el presidente norteamericano Bush y reformulada en 1994 para dirigir la política del presidente Clinton (tomado de Jennings, 1998):

OBJETIVOS	*INDICADORES*
I. Preparados para aprender	1. Indice de salud de los niños 2. Inmunizaciones 3. Lectura y explicación de cuentos por la familia 4. Asistencia a la escuela infantil
II. Completar la escolarización	5. Bachillerato finalizado
III. Rendimiento y ciudadanía de los estudiantes	6. Rendimiento en matemáticas 7. Rendimiento en lectura
IV. Formación y desarrollo profesional del profesorado	Se desarrollarán en el futuro
V. Matemáticas y Ciencia	8. Comparaciones internacionales de rendimiento en matemáticas 9. Comparaciones internacionales de rendimiento en ciencias
VI. Educación de adultos y formación permanente	10. Alfabetización de adultos 11. Participación en la educación de adultos 12. Participación en la educación superior
VII. Seguridad, disciplina y ausencia de drogas y alcohol en los centros educativos	13. Consumo de drogas y alcohol por parte de los estudiantes 14. Venta de droga en la escuela 15. Agresiones de alumnos y profesores 16. Interrupciones en clase por parte de los alumnos
VIII. Participación de los padres	Se desarrollarán en el futuro

5.3. Características de los indicadores

Álvaro (1993) resume las características de los indicadores en torno a su contenido y su calidad técnica:

- De CONTENIDO: *los indicadores educativos deben informar sobre el funcionamiento del sistema educativo en relación con las metas perseguidas, tanto de condiciones como de productos, así como sobre aquellas características que, según la bibliografía existente, influyen en los logros educativos. Estas características deben ser muy importantes, centrales y, además, de relevancia especial bajo el punto de vista político. Pues no debe olvidarse que la finalidad esencial del uso de indicadores es proporcionar información, especialmente a quien tiene poder de decisión política.*
- TÉCNICAS: *los indicadores deben medir características referidas a todo el sistema y que son perdurables en el tiempo, aunque susceptibles de ser modificadas por decisiones políticas. La información recogida debe ser comprensible para quien vaya a hacer uso de ella, así como medible, fiable y válida.*

5.4. Selección y uso de los indicadores

La selección de los indicadores debe tener en cuenta, por un lado, la demanda de los usuarios, y de otra, la disponibilidad de datos. Para permitir la comparabilidad y el estudio de la evolución del sistema, es necesario utilizar los mismos indicadores cada año y garantizar formas de recogida similares. A la hora de seleccionar los indicadores hay que tener en cuenta que se debe disponer del mismo tipo de información en varios momentos y/o de varias regiones (caso de ser nacional) o de varios países o regiones mundiales (caso de ser internacional). A estos indicadores, en la medida de que se disponga de nuevos datos o más detallados, o por su pertinencia para el tema de estudio, se añadirán otros nuevos, precisamente, para aumentar o mejorar la información. En cualquier caso un sistema de indicadores debe planificarse a largo plazo para garantizar la comparabilidad, la validez y la utilidad de la información que producen.

Blank (1991) señala los siguientes puntos a la hora de proceder al desarrollo de un sistema de indicadores educativos internacionales:

- Centrarse en la comparabilidad y especificación de los datos, basándose en tres puntos: a) consensuar los indicadores deseados; b) especificar cómo medir cada indicador y determinar si existe ya, como dato recogido, en cada país y c) decidir la opción metodológica de recogida de datos (utilizar el mismo instrumento, criterios comunes o información descriptiva).
- Garantizar la continuidad del trabajo mediante dos principios: a) asegurar el compromiso político y b) determinar el coste de cada participante. Esto llevaría al nombramiento de un comité técnico encargado de diseñar el sistema y valorar la importancia de cada indicador en base a tres criterios: a) importancia/utilidad, b) calidad técnica y c) factibilidad.
- Estandarizar la recogida de datos, el análisis y la interpretación y los informes. En este punto entraría todo aquello que facilite la comparabilidad de los datos. Así, es necesario determinar la unidad de análisis, utilizar los mismos métodos de muestreo y recogida de datos, utilizar un sistema de tabulación y análisis centralizado y diseñar un sistema de informe común (por ejemplo, determinando si la presentación de los datos se hará mediante los totales de cada estado o desagregados por grupos demográficos o desagregados por niveles educativos).

La interpretación de los indicadores es tan importante como su construcción. Un mal uso o utilización sesgada puede suponer la inutilidad del esfuerzo realizado. Un sistema de indicadores es parte de la evaluación, pero ésta es, como diría Landsheere (1996: 47-48) *algo más que acumular indicadores. La conducción tiene tres componentes necesarios: la recogida regular de información, la evaluación de dicha información y, sobre todo, la traducción en acciones institucionales o en sanciones.*

5.5. Proyectos internacionales de indicadores educativos

En los últimos años se está haciendo un gran esfuerzo por consensuar un grupo de indicadores que ofrezcan una mejor información a los responsables de la política educativa, proporcionando una mayor comprensión de la organización y funcionamiento de los sistemas educativos propios y a través de la comparación con los sistemas educativos de otros países cercanos o similares. Entre los proyectos internacionales podemos citar el *Proyecto INES* y el *Proyecto PISA*.

El *Proyecto Internacional de Indicadores de la Educación* (Proyecto INES) coordinado por el Centro para la Investigación e Innovación Educativa (CERI) de la OCDE intenta desarrollar un sistema de indicadores para que sea utilizado por los países miembros de la Unión Europea con la intención de poder comparar sus respectivos sistemas educativos así como observar la evolución en función de los cambios de la política educativa emprendida.

La primera fase del proyecto se inició en 1988. En 1991 se estableció un mecanismo consultivo internacional, a través de un Grupo Técnico y una serie de Redes de Trabajo en cada país, encargado de formular un marco común de definiciones, medidas y organización en factores educativos de acuerdo con las prioridades políticas de los países miembros. La información obtenida se recoge anualmente, desde 1992, en una publicación titulada *Education at a Glance (Análisis del Panorama Educativo).* El objetivo de esta publicación es ofrecer una visión panorámica del funcionamiento de los sistemas educativos, pudiendo tener una información del sistema propio comparándolo con el resto y, así, identificar sus puntos fuertes y débiles, ayudando a tomar las decisiones políticas más convenientes para su mejora o mantenimiento del nivel alcanzado. Los indicadores se agrupan en torno a varios factores para darle una mayor cohesión y facilitar la comprensión, como puede verse en el gráfico de la página siguiente.

El Proyecto Internacional para la Producción de Indicadores de Rendimiento de los Alumnos (Proyecto PISA) de la OCDE tiene como finalidad analizar, periódicamente y con carácter permanente, los resultados escolares (conocimientos, destrezas, competencias) de los estudiantes de 15 años de más de 30 países de todo el mundo. Cada país participante seleccionará una muestra de 150 escuelas y 4500 escolares, realizándose la evaluación en las áreas de Lectura (en la lengua dominante de instrucción), Matemáticas y Ciencias. Cada área de conocimiento se evaluará cada tres años a partir del año 2000, iniciándose con la prueba de lectura. El sistema de indicadores utilizado se revisará cada nueve años. El proyecto contempla relacionar el rendimiento con indicadores contextuales para comparar los sistemas educativos y, posteriormente, tomar decisiones de carácter político. Los indicadores contextuales que se proyecta incluir contemplarán las diferencias entre: los países; los sistemas educativos y la organización de los centros; los currí-

Costes, Recursos y Procesos Escolares

Recursos financieros.
Gastos de educación
F01: Gastos de educación en relación al PIB
F02: Gastos de las instituciones de enseñanza públicas y privadas.
F03: Gastos de enseñanza por alumno
F04: Asignación de fondos por nivel de enseñanza.
F05: Gastos de funcionamiento y gastos en capital
Fuentes de fondos educativos
F11: Fondos públicos y fondos privados
F12: Financiación pública por nivel administrativo
F13: Participación de la enseñanza en el gasto público.
Participación en la educación
P01: Escolarización en la enseñanza formal
P02: Educación preescolar
P03: Participación en la enseñanza secundaria
P04: Características de la transición de la enseñanza secundaria a la superior
P05: Acceso a la enseñanza superior
P06. Participación en la enseñanza superior
P08: Enseñanza y formación continuas para adultos

Procesos y personal.
Tiempo de enseñanza
P11: Tiempo de enseñanza por materia
P12: Horas de enseñanza
Procesos escolares
P21: Reparto de los alumnos por grupos en las clases
Recursos humanos
P31: Personal empleado en la enseñanza
P32: Ratio de alumnos a personal docente
P33: Tiempo de enseñanza
P34: Formación de los profesores
P35: Remuneración de los profesores
P36: Características de los profesores
I+D educativo
P41: Personal en I+D educativo
P42: Gastos en I+D educativo

El Contexto de la Enseñanza

Contexto demográfico
C01: Nivel general de formación de la población
C02: Nivel de formación por sexo
C03: Jóvenes y conjunto de la población
Contexto económico y social
C11: Tasa de actividad y nivel de formación
C12: Paro de los jóvenes y de los adultos
C13: Renta per cápita nacional
Opiniones y esperanzas
C21: Importancia de las materias de enseñanza
C22: Importancia de las cualidades/aptitudes
C23: Confianza pública en la escuela
C24: Responsabilidades educativas de la escuela
C25: Respeto por los profesores
C26: Prioridades en la práctica de la escuela
C27: Toma de decisiones a nivel de la escuela

Resultados de la Enseñanza

A nivel de los alumnos
R04: Progreso en lectura
R05: Frecuencia de lectura
A nivel del sistema de enseñanza
R11: Graduación en la enseñanza secundaria de segundo ciclo
R12: Graduación universitaria
R14: Títulos universitarios
R15: Personal en ciencias e ingeniería
A nivel del mercado de trabajo
R21: Paro y nivel de formación
R22: Nivel de formación y salarios
R23: Nivel de formación de los trabajadores
R24: Situación del empleo para los que terminan su formación

Gráfico 1. Modelo de indicadores de la OCDE. (Análisis del Panorama Educativo, 1995:14)

cula y las prácticas de enseñanza; la estructura y la organización de los centros educativos (tamaño del centro, recursos, etc.); la capacidad de las escuelas para paliar o reforzar las relaciones entre rendimiento y nivel socioeconómico o género.

ACTIVIDAD: ***Análisis del panorama Educativo*** (1995) pp. 167-170. Indaga: ¿De qué indicador se trata? ¿Cómo se calcula? ¿En qué países los niños reciben un mayor número de horas de enseñanza? ¿Y en qué país menos? ¿Reciben el mismo número de horas los alumnos de 14 años que los de 9? ¿Se puede afirmar que los países que ofrecen más horas obtienen mayor rendimiento?

6. LA EVALUACIÓN DEL SISTEMA EDUCATIVO EN ESPAÑA

Las evaluaciones realizadas sobre el sistema educativo en España podríamos resumirlas en tres grupos, en función del formato utilizado y agencia de referencia. El primer grupo contemplaría las estadísticas básicas relacionadas con inputs y resultados del sistema educativo. El segundo grupo, los informes demandados por los responsables políticos a diferentes comisiones o agentes, elaborados a partir de un análisis reflexivo o mediante entrevistas a diversas audiencias. Y el tercero, agruparía los planes sistemáticos de evaluación del sistema tanto la evaluación realizada en el período de reforma educativa como las evaluaciones del sistema educativo vigente.

En la actualidad serían varios los agentes de evaluación del sistema educativo, aunque cada uno de ellos aportaría información complementaria: el Instituto Nacional de Estadística, el Consejo Escolar del Estado, la Inspección Educativa, el Centro de Investigación y Documentación Educativa (CIDE), el Instituto Nacional de Calidad y Evaluación (INCE) y los institutos de calidad u organismos de las diferentes administraciones autonómicas.

Sin ánimo de ser exhaustivo se recoge, a continuación, algunos de los informes más representativos con una visión más histórica que actual, intentando poner de manifiesto que la evaluación ha estado presente, aunque utilizara diferentes procedimientos, en los cambios y valoraciones del sistema educativo español. Sin embargo, para una visión más completa debería consultarse las respectivas páginas www.

6.1. Información estadística del sistema educativo

Anualmente se publica la estadística de la Enseñanza en España elaborada por el Instituto Nacional de Estadística con el objetivo de facilitar una primera información estadística básica sobre el estado y progreso de la educación a nivel nacional, centrada en la enseñanza de: Educación Infantil, Educación Primaria, Educación Especial, Educación Secundaria Obligatoria, Bachiller, Ciclos Formativos. Se recogen datos en cada ámbito señalado de: alumnado matriculado, profesorado, centros docentes, unidades escolares, desagregadas por titularidad del centro y por comunidad autónoma.

6.2. Informes sobre el sistema educativo

Manuel Puelles (cit.en INCE, 1998) relata la evolución de la evaluación de la educación demandada por los responsables políticos. Utilizaremos este documento para resumir los informes más relevantes realizados sobre el sistema educativo español:

- El libro blanco pre Ley General de Educación (LGE) de 1970. Primer informe crítico del sistema educativo que resalta las deficiencias educativas observadas en la sociedad española, como son: las bajas tasas de escolarización en los distintos niveles educativos; un sesgo asociado a la clase social provocado por la diferenciación educativa realizada a partir de los 10 años; la escasa valoración social del profesorado; y otros temas estructurales (como la ratio) o de contenido (memorización), etc. La LGE intentó paliar algunos de estos problemas y los desajustes fueron corrigiéndose con medidas legislativas posteriores.
- El informe *que eleva la Comisión Evaluadora de la LGE y Financiamiento de la Reforma Educativa en cumplimiento del Decreto 186/76, de 6 de febrero*, encargado por el ministro Robles Piquer a una comisión, trataba de evaluar la aplicación de la LGE, valorando sus logros y deficiencias después de seis años de su puesta en marcha. Dicho informe destacaba como puntos fuertes la casi generalización de la EGB hasta los 14 años; el aumento del alumnado en la enseñanza secundaria; el crecimiento de la enseñanza pública y la orientación "unificada y polivalente" del Bachillerato. Como puntos débiles: la no gratuidad de la enseñanza obligatoria, las deficiencias físicas de aulas y centros escolares, la elevada ratio alumno-profesor, la saturación de materias de los planes de estudios en secundaria superior, la falta de formación del profesorado de secundaria y la discriminación y tratamiento de la formación profesional, todo ello unido a una escasa financiación. Por primera vez, se habla en términos de calidad de la enseñanza y en la *Síntesis actualizada del III Informe Foessa (1978)* se compara el gasto en educación con los países de la entonces OCDE (Organización de Cooperación y Desarrollo Económico).
- *Informe sobre las Enseñanzas Medias en España* (1981) donde se evalúa el Bachillerato y la Formación Profesional, destacando sus deficiencias: el alto número de alumnos que no consiguen el título de Graduado Escolar al finalizar la EGB, el alto nivel de abandono de la FP, un alto índice de alumnos que acceden a la FP sin suficiente preparación, la función propedéutica del Bachillerato, o el fracaso del Curso de Orientación Universitaria. Tras ello, se finaliza pidiendo una urgente reforma.
- *Informe de la OCDE* (1986) sobre la política educativa española que pone de manifiesto la rápida expansión de la educación en la sociedad española, si bien también pone en evidencia la escasa financiación, y deficiencias en la formación del profesorado de secundaria, en la de los directores de los centros y en el desarrollo de la FP.
- *Libro Blanco para la Reforma del Sistema Educativo* (1989) revela deficiencias en la educación preescolar, un alto índice de fracasos al finalizar

la EGB, la temprana discriminación que supone el mantenimiento de una doble titulación al finalizar la escolaridad obligatoria, el fracaso de la FP y la ausencia de un plan de estudios secuenciado para toda la escolaridad. Por otra parte, se manifiesta la necesidad de ajustar el sistema educativo a las nuevas exigencias político-sociales de España y de los países de la Comunidad Europea.

- *Centros educativos y calidad de la enseñanza: propuesta de actuación* (MEC, 1994) se analizan algunos problemas detectados en la aplicación de la LOGSE relacionados, básicamente, con la organización y los recursos de los centros educativos, como: deficiencias en la formación, funciones y selección de los directores, baja motivación y formación en ejercicio del profesorado, limitada autonomía de los centros en la gestión de sus recursos (personales, económicos y administrativos), falta de procedimientos de evaluación de los diferentes componentes del sistema, deficiencias en la selección y formación de los inspectores. A partir de estas observaciones se proponen 77 medidas para mejorar la calidad del sistema educativo cuya viabilidad se cuestiona en el V Informe Foessa además de poner en evidencia la disminución del gasto en educación.
- *Informes Anuales del Consejo Escolar del Estado* sobre el Sistema Educativo, recogen las opiniones que sobre el funcionamiento del sistema educativo tienen las diferentes audiencias (padres, alumnos, sindicatos, patronal, administración). El Consejo Escolar del Estado fue creado por la Ley Orgánica del Derecho a la Educación (LODE 8/1985, art.30) como un órgano de ámbito nacional para la participación de los sectores afectados en la programación general de la enseñanza y para asesorar los proyectos de ley o reglamentos propuestos o dictados por el gobierno. Emite, desde mediados de los años ochenta, un informe anual sobre la situación del sistema educativo. Sus informes tienen carácter descriptivo (no valorativo) e incluyen recomendaciones a los responsables políticos. Sin embargo, estas recomendaciones no tienen valor prescriptivo, de hecho una de sus quejas es que no se atienden sus recomendaciones. Entre los problemas que describen con mayor frecuencia podemos citar: el fracaso escolar; el alto número de profesores con destino provisional y la consiguiente baja consolidación de los equipos; la desmotivación del profesorado; el funcionamiento irregular de los órganos de participación; la escasa atención a la enseñanza de lenguas extranjeras, a la educación intercultural, a la educación física y a la educación musical; el escaso gasto público dedicado a la educación, inferior al de los países de la Unión Europea; la insatisfacción con el servicio de Inspección Técnica; la escasa participación de la comunidad educativa; el progresivo deterioro del clima de convivencia en los centros escolares; etc.

- *Informes Anuales de la Inspección de Educación* sobre los centros, los resultados escolares, los programas educativos, los planes de estudio, etc. El art. 61.2.b de la LOGSE señala como una de las funciones de la Inspección, participar en la evaluación del sistema educativo, y el art. 30 de la LOPEG concreta dicha participación en "lo que corresponde a los centros escolares, a la función directiva y a la función docente, a través del análisis de la organización, funcionamiento y resultados de los mismos". Sería la evaluación del sistema a nivel mesoeducativo o institucional (centros y programas). Así, desde el curso 1986-87, se publica anualmente un informe sobre los resultados de las evaluaciones de los alumnos de Primaria y Secundaria, presentando los datos en porcentajes de alumnos que promocionan a partir de las calificaciones de los profesores. En estos informes se constata que las áreas con un índice mayor de suspensos son Matemáticas, Lengua Española, Lengua Extranjera (inglés) y Ciencias de la Naturaleza. Junto a la descripción de los resultados se incluye un análisis de los posibles factores de incidencia (actitud hacia la asignatura, programas, metodología, horarios...). Además de estos informes centrados en los resultados de los alumnos, la Inspección elabora periódicamente otros informes sobre aspectos particulares de los centros educativos como el del "funcionamiento de los Consejos Escolares" y participa en proyectos de evaluación, como la realizada sobre la "organización y funcionamiento de los centros públicos", también conocido como Plan EVA (informe publicado en 1995 por la Subdirección General de Inspección Educativa. MEC).

ACTIVIDAD: Subraya los puntos débiles observados en los informes de evaluación comentados y registra qué se ha modificado en el sistema educativo actual y qué permanece todavía por resolver.

6.3. Planes de investigación evaluativa

En este grupo se incluyen los estudios realizados sobre diferentes etapas o elementos del sistema educativo español planificados desde un punto de vista empírico. Dos son los organismos que se han ocupado de esta función: el Servicio de Evaluación del Centro de investigación y documentación educativa (C.I.D.E.) y el Instituto Nacional de Calidad y Evaluación (I.N.C.E.). Asimismo deberíamos incluir los distintos organismos que se ocupan de la evaluación a nivel de Comunidades Autónomas (ver tabla de direcciones). Hemos optado por resumir únicamente el informe pre-LOGSE e incorporar, esquemáticamente, las funciones del INCE y

sus líneas de actuación. Remitimos al lector a ampliar la información en las páginas www respectivas.

- *La Evaluación Externa de la Reforma de Enseñanzas Medias por el Centro de Investigación y Documentación Educativa (1984-89).* Este estudio se considera la primera experiencia de evaluación sistemática del sistema, en este caso, en período de reforma educativa. En palabras de A. Tiana (1990) supone un punto de inflexión importante por ser una nueva práctica de la administración educativa española. El objetivo era analizar si la enseñanza experimental conseguía logros educativos diferentes a los obtenidos por la enseñanza vigente (BUP y FP). Se elabora un plan de evaluación longitudinal con varias fases, contrastando los resultados de los alumnos en rendimiento, aptitudes, actitudes, expectativas, etc., controlando por variables que pudieran incidir en posibles diferencias. Las conclusiones se agrupan en torno a los resultados y al funcionamiento del sistema. Entre los resultados más destacados podríamos señalar resultados similares en rendimiento entre los grupos de BUP y reforma y mayores entre éstos y FP, excepto en cálculo numérico. Mejores resultados en aptitudes de los grupos de BUP. Mejores resultados en los grupos de reforma en variables actitudinales, satisfacción, y en las tasas de abandono y repetición u una mayor percepción de la utilización de una metodología activa y participativa del profesorado. Asimismo, se pone en evidencia que la reforma produce fenómenos de igualación, al reducir las diferencias existentes entre los alumnos de BUP y FP.
- Estudios del Centro de Investigación y Documentación Educativa sobre calidad y equidad:

> **ACTIVIDAD:** Consulta la página www del CIDE y describe, brevemente, los estudios que ha financiado sobre calidad y equidad del sistema educativo. Cita algún otro trabajo publicado que esté relacionado con la evaluación del sistema educativo.

- *Evaluaciones del Instituto Nacional de Calidad y Evaluación.* En el artículo 62 de la LOGSE se encomienda la evaluación del sistema educativo al Instituto Nacional de Calidad y Evaluación (INCE). El 18 de junio de 1993 se aprueba el Real Decreto 928/1993 donde se detallan las funciones, composición y atribuciones de sus órganos de gobierno y los principios básicos de su organización. El ámbito de actuación corresponde al Estado, encargándose de evaluar los aspectos del currículo que constituyen las enseñanzas mínimas establecidas en el artículo 4 de la LOGSE para las diferentes etapas, ciclos y niveles del sistema educativo.

  Entre las funciones que se le asignan, destacaríamos:

- Analizar el grado de eficacia y de eficiencia del Sistema Educativo.
- Elaborar sistemas de evaluación para las diferentes enseñanzas reguladas en la LOGSE.
- Coordinar la participación española en los estudios internacionales sobre evaluación.
- Colaborar y cooperar con las Administraciones educativas en materia de evaluación.
- Proporcionar e intercambiar información con las Administraciones Educativas para la toma de decisiones.
- Proporcionar información al Consejo Escolar del Estado.
- Informar a la sociedad sobre el funcionamiento y resultados del sistema educativo.
- Difundir los resultados de sus evaluaciones e investigaciones.

Desde 1994 se han puesto en marcha acciones concretas para evaluar el sistema educativo en sus diversas etapas y en diversas áreas específicas de conocimiento (como la Evaluación de la Educación Primaria; el Diagnóstico de la Escuela Secundaria Obligatoria; Evaluación de la Educación Física en la educación primaria). Al mismo tiempo, colabora en evaluaciones internacionales (Estudio Internacional sobre la Enseñanza y el Aprendizaje de la lengua Inglesa en la Educación Secundaria; Programa de Evaluación de la Calidad de la Educación de la OEI) y en la elaboración de un sistema de indicadores internacional para evaluar la calidad educativa (Proyecto INES -Proyecto Internacional de Indicadores de la Educación del Centro para la Investigación e Innovación Educativa (CERI)- o en el Proyecto PISA -Proyecto Internacional para la Producción de Indicadores de Resultados Educativos de los Alumnos-).

- *Evaluaciones de las Instituciones de Evaluación de las Comunidades Autónomas.* Al tiempo que la LOGSE asigna al instituto de evaluación la evaluación del sistema educativo, establece que las administraciones educativas autonómicas evalúen el sistema educativo en el ámbito de sus competencias. En la medida que las transferencias a las Comunidades Autónomas en materia educativa se ha realizado, en muchas de ellas se han creado instituciones específicas encargadas de la evaluación general del sistema en la comunidad de referencia. Sus funciones son similares al INCE aunque se hacen específicas para responder a las necesidades y demandas contextuales añadiendo la de colaboración y participación en las evaluaciones de carácter estatal o internacional. En cada caso se cuenta con diferente publicaciones referidas a las evaluaciones realizadas, para facilitar el acceso a dicha información, incluimos los organismos que, en cada Comunidad Autónoma, se encargan de las funciones de la evaluación del sistema educativo.

ORGANISMO ENCARGADO DE LA EVALUACIÓN	DIRECCIÓN
INCE Instituto Nacional de Calidad y Evaluación www.ince.mec.es	C/ S. Fernando del Jarama, 14 28002 MADRID 34 91 7459200
C.A. ANDALUCÍA Dirección General de Evaluación Educativa y Formación del Profesorado	Isla de la Cartuja, s/n 41071 SEVILLA 95 4464800
C.A. ISLAS BALEARES Consejería de Educación, Cultura y Deportes	C/ S. Felió, 8-A 07012 PALMA DE MALLORCA 971 176500
C.A. CANARIAS Instituto Canario de Evaluación y Calidad Educativa http://nti.educa.rcanaria.es/icec/	C/ Pedro de Vera, 36 35003 LAS PALMAS 928 381747
GENERALITAT CATALUNYA Consell Superior d'Avaluació del Sistema Educatiu	Via Augusta, 202-226, 1ºD 08021 BARCELONA 93 4006913
C.A. GALICIA Servicio de Ordenación e Innovación Educativa	C/ San Lázaro, 107 15704 SANTIAGO DE COMPOSTELA 981 546504
COMUNIDAD FORAL NAVARRA Servicio de Inspección Técnica y Servicios	Cuesta Sto. Domingo, s/n 31001 PAMPLONA 948 426500
C.A. PAÍS VASCO Dirección de Renovación Pedagógica	C/ Duque de Wellington, 2 01010 VITORIA-GASTEIZ 945 249900
C. VALENCIANA Servicio de Evaluación e Inspección	Avd. Campanar, 32 46015 VALENCIA 96 3863217

BIBLIOGRAFÍA COMENTADA

TIANA, A. y SANTANGELO, H. (1996). Evaluación de la calidad de la educación. *Revista Iberoamericana de Educación*, 10, 215-230

Se trata de un documento elaborado como base de discusión a la VII reunión de países iberoamericanos. Se describe brevemente el origen, las tendencias, los modelos, los instrumentos y la utilización política de la evaluación de los sistemas educativos, vinculando dos conceptos básicos actuales como son la calidad y la evaluación. Aporta una visión global sobre el tema, desde una panorámica más política que técnica pero que es fundamental por la vinculación de la evaluación de los sistemas educativos con la toma de decisiones político-administrativas.

LANDSHEERE, G. de (1996). *El pilotaje de los sistemas educativos*. Madrid: La Muralla.

Libro que intenta recoger elementos clave de la evaluación de los sistemas educativos desde experiencia del autor, vinculado a proyectos de evaluación internacional. El libro se distribuye en dos grandes bloques, uno más teórico donde se plantea la evaluación, o pilotaje, ligada a la dirección, gestión y toma de decisiones de los sistemas educativos, se describe la naturaleza, los tipos, los criterios, la metodología y los indicadores de la evaluación educativa; otro, donde se ejemplifican modelos de evaluación a nivel nacional (USA, Inglaterra, Francia, Canadá, etc.) e internacional (IEA, IAEP, OCDE).

SLEDSENS, G.; VANDEWIELE, M. y WOLFS, J.L. (1996). L'évaluation des systèmes éducatifs. *Scientia Paedagogica Experimentalis*, XXXIII, 2, 181-200.

Desarrolla desde una perspectiva práctica los elementos clave para poner en práctica una evaluación del sistema educativo, denominada macroevaluación por los autores. De esta forma responde, brevemente, a cuestiones relativas al objeto de evaluación, los criterios, las audiencias, las fases, la metodología y los informes.

ÁLVARO, M. (1993). Los indicadores de calidad de educación. *Revista de Psicología General y Aplicada*, 46 (2) 221-232.

Documento básico sobre los indicadores educativos que plantea el origen y los motivos que han llevado a desarrollar y utilizar indicadores en educación, asimismo describe el concepto, características y tipología. Por último detalla los problemas más frecuentes en el establecimiento de sistemas de indicadores y sintetiza una serie de recomendaciones útiles para su desarrollo y utilización con el fin de permitir una comparación eficaz y funcional entre diferentes sistemas educativos o diferentes momentos.

DIRECCIONES ÚTILES EN RELACIÓN CON LA EVALUACIÓN DEL SISTEMA EDUCATIVO

ORGANISMO	DIRECCIÓN	INFORMACIÓN PROPORCIONADA
INCE Instituto Nacional de Calidad y Evaluación www.ince.mec.es	ESPAÑA C/ S.Fernando del Jarama, 14 28002. Madrid 34 91 7459200	• Evaluación del Sistema Educativo Español. • SIE (Sistema de Indicadores de Educación) Elaboración del Sistema nacional de Indicadores (valorar eficacia y eficiencia). • Participación en proyectos de evaluación internacionales.
Instituto Nacional de Estadística www.ine.es www.mec.es/estadística	ESPAÑA	• Estadísticas del INE (estadísticas del INE: *España en cifras* (datos demográficos, sociales y económicos). • Anuario de Estadística Universitaria (Consejo de Universidades). • Oficina de Planificación y Estadística del MEC en colaboración con las Comunidades Autónomas.
Consejo Escolar del Estado www.mec.es	ESPAÑA	• Informes anuales sobre el estado y situación del sistema educativo.
OCDE CERI Centro para la Investigación e Innovación Educativa	2 rue André Pascal 75775 París FRANCIA	• Análisis del Panorama Educativo. • Estadísticas de Educación de la OCDE (incluye los datos primeros para calcular los indicadores y las series históricas a partir de 1985). • Proyectos de Indicadores. • International Adult Literacy Study.
EUROSTAT Oficina de estadística de las Comunidades Europeas europa.eu.int/en/comm/eurostat	Comunidad Europea	• Proporcionar información estadística de apoyo a las instituciones europeas para establecimiento de políticas comunitarias. • Elaborar estadísticas de interés general. • Establecer sistema común que enlace los sistemas estadísticos de los Estados miembros. • Cooperar técnicamente con otros organismos estadísticos del resto del mundo. • Revista mensual Just published. • Boletín trimestral Sigma.
EURYDICE Red Europea de Educación (compuesta por los Ministerios de Educación de los países participantes y la Comisión Europea). Creada en 1980 www.eurydice.org	Comunidad Europea	• Base de datos Eurybase: descripción de los sistemas educativos nacionales, con estructura común para facilitar análisis comparativos. Se incluye legislación, instituciones, bibliografía, etc. • European Education thesarurus (diccionario europeo de sinónimos de la educación en los 11 idiomas oficiales. Formato PDF). • Comparative Studies: publicaciones, informes y análisis sobre la educación.

ORGANISMO	DIRECCIÓN	INFORMACIÓN PROPORCIONADA
IEA International association for the Evaluation of Eductional Achievement Centros de investigación de más de 50 países 1958 http://www.iea..nl	Internacional	• Evaluación periódica de los resultados logrados por los alumnos (matemáticas, ciencias, sociales, lectura, lenguas extranjeras, educación cívica...). • Análisis comparativos transnacionales a partir de las diferencias existentes entre países (profesorado, estructura escolar, currículo, duración del curso académico, tamaño de la clase....). • Análisis longitudinales de los logros en diferentes áreas (matemáticas y ciencias).
NAEP National Assessment of Educational Progress nces.ed.gov/nationsreportcard/	EEUU 1860,Lincoln Street Denver. Colorado.	• Informes de las evaluaciones (a partir de 1969-70). • Acceso a los resultados de las evaluaciones con fines de estudio.
SINEC Sistema Nacional de Evaluación de la Calidad	ARGENTINA	• Evaluación del sistema educativo argentino y ofrecer políticas de mejora de la calidad (desde 1993).
DEP Dirección de la Evaluación y de la Prospectiva. Ministerio de Educación Nacional	FRANCIA	• L'état de l'école (anual, desde 1991). Publicación que recoge 30 indicadores sobre su sistema y su evolución en costes, actividades y resultados en primaria, secundaria, superior y formación continua. • Géographie de l'école (anual desde 1993) describe las características por regiones y distritos educativos. Incluye: entorno social y cultural, recursos y utilización, funcionamiento, aprendizaje, inserción.
AEI Assessment Effectiveness Innovation http://uttou2.to.utwente.nl/AEI/	Internacional	• Resultados estadísticos.
Ministerio de Educación e Investigación Danés	DINAMARCA	• Sistema de 40 indicadores agrupados en recursos, actividades y resultados.
Ministerio de Educación del Gobierno de Quebec	CANADA	• Sistema de indicadores sobre educación primaria y secundaria agrupados en: esfuerzo financiero, trayectoria escolar, evaluación de los aprendizajes, diplomas de secundaria, educación de adultos.

BIBLIOGRAFÍA GENERAL

ABRAMI, P.C. (1989). "How should we use student ratings to evaluate teaching", en *Research in Higher Education,* 30 (2); pp. 221-227.

ACHERMAN, H. (1992). Evaluación de la calidad por el profesorado. *Actas del Congreso Internacional de Universidades.*

AGUILAR, Mª.J. y ANDER-EGG, E. (1992). *Evaluación de servicios y programas sociales*. Madrid: Siglo XXI.

ALKIN, M.C. (1969). "Evaluation Theory development", en *Evaluation Comment,* 2, págs. 2-7.

ALLEN, R. et al. (1990). *The Geography learning of high School Seniors.* New Jersey: NAEP.

ÁLVAREZ, M. (1998). *El liderazgo de la calidad total.* Madrid. Escuela Española.

ÁLVAREZ MÉNDEZ, J. (1992). "La ética de la calidad", en *Cuadernos de Pedagogía*. 199, Enero. 8-12.

ÁLVAREZ MOLINERO, L. (1989). "Evaluación de la docencia en la U.A.M.", *en Boletín del Instituto de Ciencias de la Educación*, Madrid, n. 14: p. 35-37.

ÁLVAREZ ROJO, V. (Dir.) (2000). *Propuestas del profesorado bien evaluado para potenciar el aprendizaje de los estudiantes.* Vicerrectorado de Calidad. Sevilla. Instituto de Ciencias de la Educación de la Universidad de Sevilla.

ÁLVARO, M. (1993). *Elementos de psicometría*. Madrid: Eudema.

ÁLVARO, M. (1993). "Los indicadores de calidad de educación", en *Revista de Psicología General y Aplicada*, 46, 2, 221-232.

ÁLVARO, M. et al. (1988). *Evaluación externa de la Reforma Experimental de las Enseñanzas Medias (I).*Madrid: CIDE.

ÁLVARO, M. et al. (1990). *Evaluación externa de la Reforma Experimental de las Enseñanzas Medias (II)*. Madrid: CIDE.

ÁLVARO, M. et al. (1992). *Evaluación externa de la Reforma Experimental de las Enseñanzas Medias (III)*. Madrid: CIDE.

ALVIRA MARTÍN, F. (1991) *Metodología de la evaluación de programas*. Madrid: Centro de Investigaciones Sociológicas.

AMADOR MUÑOZ, L. y DOMÍNGUEZ LEÓN, J. (Coord.) (1996). *Mejorar cada día. Evaluación y calidad de la enseñanza*. Uned. Centro Asociado de Sevilla. Junta de Andalucía. Consejería de Educación y Ciencia.

APODACA, P. y LOBATO, C. (Eds.) (1997). *Calidad en la universidad: orientación y evaluación*. Barcelona: Laertes.

APODAKA, P. (1990). "Experiencias evaluativas en la Universidad del País Vasco: consideraciones en torno a puesta en marcha de un proceso", en *Revista Española de Pedagogía*, Madrid, n. 186; p. 327-336.

ARIAS, B.; VERDUGO, M.A. y RUBIO, V. (1995). *Evaluación de la calidad y modelo local de Valladolid* (Programa Helios). Ministerio de Educación y Ciencia.

ARNAL, J.; DEL RINCÓN, D. y LATORRE, A. (1992). *Investigación educativa. Fundamentos y metodología*. Barcelona: Labor.

ARROYO SIMÓN, M. (1990). "Evaluación de la calidad de la educación norteamericana", en *Revista Complutense de Educación*, 1, 2, 207-221.

ASTIN, A.W. (1985). *Achieving Educational Excellence*. San Francisco: Jossey-Bass.

BALCELLS, J. (1994). *La Investigación Social*. Barcelona: PPU.

BALL, R. y HALWACHI, J. (1987). "Perfomance Indicators in Higher Education", en *Higher Education*, 16, 393-405.

BALLES, M.S. (1997). *Técnicas cualitativas de investigación social*. Madrid: Síntesis.

BARROSO, M.C. (1988). "Progreso y Calidad en Educación". En *La Calidad de los Centros Educativos*. Actas del IX Congreso Nacional de Pedagogía. Alicante.

BEARE, H. y otros (1992). *Cómo conseguir centros de calidad*. Madrid: La Muralla.

BEATON, A.E., et al. (1996). *Mathematics achievement in the middle school years: IEA's Third International Mathematics and Science Study*. Chesnut Hill: Boston College.

BEATON, A.E., et al. (1996). *Science achievement in the middle school years: IEA's Third International Mathematics and Science Study*: Chesnut Hill. Boston College.

BENEDITO, V. (1989). "La evaluación del profesor universitario", en *Revista de Educación*, Madrid, n. 290; p. 279-291.

BERBEROGLU, G.; DOCHY, F. y MOERKERKE, G. (1996). Psychometric evaluation of entry assessment in higher education: A case study. En *European Journal of psycology of Education*, XI, 1: 15-43.

BERNILLON, A. y CERRUTTI, O. (1989). *Implantar y Gestionar la Calidad Total*. Barcelona: Ediciones Gestión 2000, S.A.

BLANCO, A. (1993). "Fiabilidad, precisión, validez y generalización de los diseños observacionales". En ANGUERA, M.T. (Ed). *Metodología observacional en la Investigación Psicológica*. Barcelona: PPU.

BLANCO, A. y SENLLE, A. (1988). *Calidad Total. Desarrollo de directivos y calidad total en la organización*. Madrid: ESIC.

BLANCO FELIP, L. (1993). "Modelo para una autoevaluación modular en una institución educativa (B.A.D.I.)", en *Revista Investigación Educativa*, Barcelona, n. 21; p. 47-66.

BLANK, E.K. (1991). "Developing and implementing educational indicators". *Ponencia presentada al Seminario de Evaluación de Sistemas e Indicadores Educativos*. Madrid: CIDE.

BLAT, J. (1994). *Primer Congreso Nacional sobre nuevas Tecnologías Aplicadas a la Formación*. Madrid.

BORRERO CABAL, A. (1992). "Evaluación institucional universitaria: el nivel académico". *Actas del Congreso Internacional de Universidades*. Madrid.

BRANDON, P.R.; LINDBERG, M.A. & WANG, Z. (1993). "Involving Program Beneficiaries in the Early Stage of Evaluation: Issues of Consequentiel Validity and Influence". *Educational Evaluation and Policy Analysis*. Vol. 15, nº 4, pp. 420-428.

BRICALL, J.M. (1998). "Participación de la sociedad en las universidades". En PORTA, J. y LLADONOSA, M. (Coords.). *La universidad en el cambio de siglo*. Madrid: Alianza Editorial.

BRINKERHOFF, R.O.; BRETHOWER, D.M.; HLUCHYJ, T. y NOWAKOWSKI, J.R. (1985). *Program Evaluation. A Practitioner's Guide for Trainers and Educators*. Boston: Kluwer-Nijhoff Publishing.

BROOKOVER, W. et al. (1979). *School social systems and student achievement Schools can make a difference*. New York: Praeger Publishers.

BUENDÍA, L. (1999). *La preocupación de la universidad por la calidad*. Huelva. Inédito.

BUENDÍA, L. y PEGALAGAR, M. (1990). "La Teoría de la Generalizabilidad como base para la evaluación y toma de decisiones", en *Revista de Investigación educativa*, 8, 16: 483-491.

BUENDÍA EISMAN, L.; COLÁS BRAVO, P. y HERNÁNDEZ PINA, F. (1997). *Métodos de Investigación en Psicopedagogía*. Madrid: McGraw Hill.

BUNDERSON, V.; DILLON, K. y OLSEN, J. (1993). "The Four Generations of Cumputerized Educational Measurement". En LINN, R. (Ed.). (1993). *Educational Measurement*. New York: Macmillan Pub. Co.

CABRERA, L.; GOBANTES, J.M.; GONZÁLEZ, S. y PÉREZ SANJUÁN, N. (2000). *Plan Municipal de Educación y Atención a la Sexualidad Juvenil del Municipio de Los Realejos, Tenerife. INFORME DE EVALUACIÓN*. Documento multicopiado. Enero de 2000.

CAJIDE, J. y TEJEDOR, J. (1988). "Modelos de funcionamiento institucional y calidad de los centros educativos". *Actas del IV Seminario de modelos de Investigación Educativa*. Santiago de Compostela.

CANAL DÍEZ, J.A. (1988). "Evaluar la función docente", en *Andecha Pedagogica*, Gijón, n. 19; p. 8-12.

CANO GARCÍA, E. (1998). *Evaluación de la Calidad Educativa*. Madrid: La Muralla.

CARDONA ANDÚJAR, J. (Coord.) (1994). *Metodología innovadora de evaluación de centros educativos*. Madrid: Sanz y Torres.

CEA D'ANCONA, M.A. (1996). *Metodología cuantitativa. Estrategias y técnicas de investigación social*. Madrid: Síntesis.

CHORDI CORBO, A. (1988). "Evaluación periódica del profesorado y evaluación de la eficacia del aprendizaje", en *Studia Paedagogica*, Salamanca, n. 20; p. 35-40.

CLARKE, D. y STEPHENS, M. (1996). "The Ripple Effect: The Instructional Impact of the Sistematic Introduction of Performance Assessment in Mathematics". En BIRENBAUM, M. y DOCHY, F. (Ed.). *Alternatives in Assessment of Achievements Learning Process and Prior Knowledge*. Boston: Kluwer Academic.

COBO, J.M. (1985). "El reto de la calidad en la educación. Propuesta de un modelo sistemático", en *Revista de Educación*, nº 308, pp. 358.

COLÁS, P. y BUENDÍA, L. (1992). *Investigación Educativa*. Sevilla: Alfar

COLÁS BRAVO, Mª.P. y REBOLLO CATALÁN, Mª.A. (1993). *Evaluación de Programas. Una guía práctica*. Sevilla: Kronos.

COLEMAN, J.S. (1966). *Equality of educational opportunity*. Washington, D.C.: U.S. Government Printing Office.

COMISIÓN EUROPEA (1997). *Las cifras clave de la educación en la Unión Europea*. Luxemburgo: Publicaciones de la C.E.

CONSEJO DE UNIVERSIDADES (1988). *Plan de Evaluación de la Calidad de las Universidades. Guía de Evaluación*. Madrid.

CONSEJO DE UNIVERSIDADES (1993). *Programa Experimental de la Calidad del Sistema Universitario*. Madrid: Consejo de Universidades. Secretaría General.

CONSEJO DE UNIVERSIDADES (1998). *Plan Nacional de Evaluación de la Calidad de las Universidades. Guía de Evaluación*. Madrid: Consejo de Universidades. Secretaría General.

CONSEJO ESCOLAR DEL ESTADO (1999). *Informe sobre el estado y situación del sistema educativo. Curso 1997-98*. Madrid: MEC.

COOK, S.S. (1989). "Improving the quality of student ratings of instruction: A look at two strategies", en *Research in Higher Education*, 30 (1); pp. 31-45.

CORTÉS, J. (1995). "El proceso de implantación de la TQM", en *Organización y Gestión Educativa*, 3, pp37-40.

CORTINA, A. (1997). *Ciudadanos del mundo. Hacia una teoría de la ciudadanía*. Madrid: Alianza Editorial.

COWAN, J. (1985). "Effectiveness and efficiency in higher education", en *Higher Education*, 14, 235-239.

CRANSTON, N. (1991). "IBSO/A proyect". En HEWTON, J. (Edit.). *Performance indicators in education*. Australia Brinsbane: ACDGE.

CREEMERS, B. y SCHEERENS, J. (1989). "Conceptualizing school effectiveness", en *Intemational Journal of Educational Research*, 13 (7), 691-706.

CRONBACH, L.J. (1963). "Course improvement throught evaluation", en *Teachers college Record*, 64, págs. 672-683.

CRONBACH, L.J. (1972) (2ª ed.). *Fundamentos de la exploración psicológica*. Madrid: Biblioteca Nueva.

CRONBACH, L.J. (1980). *Toward reform of program evaluation*. San Francisco: Jossey-Bass.

DARDER, P. y LÓPEZ, J.A. (1989). *QUAFE-80. Cuestionario para el análisis del funcionamiento de la escuela*. Barcelona: ONDA.

DE LA ORDEN, A. (1993) "La escuela en la perspectiva del producto educativo. Reflexiones sobre evaluación de centros educativos", en *Bordón*, Vol. 45, nº 3, pp.263-270.

DE LA ORDEN, A. (1995a). "Evaluación de la organización de centros escolares", en *Revista Galega de Psicopedagoxía*. Monográfico nº 6, 51-66.

DE LA ORDEN, A (1995b). "Hacia un modelo sistémico para la evaluación de la calidad universitaria", en *Revista Galega de Psicopedagoxía*. Monográfico nº 6, 147-161.

DE LA ORDEN, A. et al. (1997). "Desarrollo y validación de un modelo de calidad universitaria como base para su evaluación", en *Revista Electrónica de Investigación y Evaluación Educativa*, 1-2, (3).

DEL RINCÓN, D.; ARNAL, J.; LATORRE, A. y SANZ, A. (1995). *Técnicas de investigación en ciencias sociales*. Madrid: Dykinson.

DE LUXÁN, J.M. (1998). "La evaluación de la universidad de España", en *Revista de Educación, 315*, 11-28.

DE MIGUEL, M. (1989). "Modelos de investigación sobre organizaciones educativas", en *Revista de Investigación Educativa*, 7, 13, 21-56.

DE MIGUEL, M. (1990). "Indicadores de calidad en la docencia universitaria". *Ponencia presentada en el Congreso sobre Calidad de la Educación universitaria*. Puerto de Santa María. Marzo.

DE MIGUEL, M. (1995). "La calidad de la educación y las variables de proceso y de producto. Donostia. Eusko Ikaskuntza", en *Cuadernos de Sección. Educación,* 8. p. 29-51.

DE MIGUEL, M.; MADRID, V.; NORIEGA, J. y RODRÍGUEZ, B. (1994). *Evaluación para la calidad de los Institutos de Educación Secundaria*. Madrid: Escuela Española.

DE MIGUEL, M.; MORA, J.G. y RODRÍGUEZ, S. (1991). *La evaluación de las instituciones universitarias.* Madrid: Consejo de Universidades. Secretaría General.

DE MIGUEL DÍAZ, M (1997). "La evaluación de los centros educativos. Una aproximación a un enfoque sistémico", en *Revista de Investigación Educativa*, Vol. 15, nº 2.

DEMING, E.W. (1981). *Management and statistical Techniques for quality and productivity*. New York: New York University.

D'HAINAUT, L. (1988). *Sistemas educativos. Análisis y regulación*. Madrid: Narcea.

DOCHY, F. (1992). *Assessment of Prior Knowledge as a Determinant for Future Learning: The Use of Prior Knowledge State Tests and Knowledge Profiles.* Utrech/London: Lemma/Jessica Kingsley Publisher.

DOCHY, F. y MOERKERKE, G. (1997). "Assessment as a mayor influence an learning and instruction", en *International Journal of Educational Research,* 27: 415-432.

DOCHY, F.J.R.C.; SEGERS, M.S.R. y WIJNEN, W.H.F.W. (1990). "Selecting Performance Indicators". En GOEDEGEBUURE, L.C.J.; MAASSEN, P.A.M. y WESTERHEIJDEN, D.F. (Eds). *Peer Review and Performance Indicators.* Utrecht, NL: Vitjeverij Lemma.

DONALDSON, J.; STAROPOLI, A.; OTTENWAELTER, M.O.; THUNE, C. y VROEIJENSTIJN, T. (1994). *Projects pilotes europeens pour l'evaluation de la qualité dans l'enseignement supérieur: lignes directrices pour les établissements participants.* Comision de las Comunidades Europeas.

DUNKIN, M.J. y BARNES, J. (1986). "Research on teching in Higher Education", en WITTROCK, M.C.: *Handbook of research on teching.* New York: MacMillan.

EBEL, R. (1997). *Fundamentos de la medición educativa.* Buenos Aires: Guadalupe.

EDMONS, R. (1979). *A discussion of literature and related to effective schooling.* Cambridge Harward Graduate School of Educational.

EEC (1991). *Memorandum on Higher Education in the European Community.* Brussels.

EEC (1993). *Conference of European Rectors. View on the Memorandum on HE en the EEC.* Brussels.

EISNER, E.W. (1975). *The perceptive eye: toward the reformation of Educatiobal evaluation.* Sanford (Calif.): Standard Evaluation Consortium.

EISNER, E.W. (1981). *The methodology of Quantitative evaluation: The case of educational Connoisseurship and educational Criticism.* Stanford: Stanford University.

ELLIOT, J. (1993). "Conocimiento, poder y evaluación del profesor". En W. CARR (1989 orig.; 1993 trad.), *Calidad de la Enseñanza e Investigación-Acción*. Sevilla. Diada Editora. pp:115-131.

ELLIOT, J. et al. (1985). *Investigación/acción en el aula.* Valencia: Generalitat Valenciana.

ESCOTET, M.A. (1984). *Técnicas de evaluación institucional en la educación superior*. Madrid: MEC.

ESCUDERO, T. (1989). "Aproximación pragmática a la evaluación de la universidad", en *Revista de investigación educativa*, vol. 7, 13, pp.93-112.

ESCUDERO, T. (1999). *Indicadores del rendimiento académico: una experiencia en la Universidad de Zaragoza.* Inédito.

ESCUDERO ESCORZA, T. (1980). *¿Se pueden evaluar centros docentes y sus profesores?* Zaragoza: I.C.E.

ESPAÑA. PLAN NACIONAL DE INVESTIGACIÓN CIENTÍFICA Y DESARROLLO TECNOLÓGICO (1996). *Instituciones y centros de investigación y desarrollo.* Madrid: Comisión Interministerial de Ciencia y Tecnología.

ESTEBAN, M.C.y MONTIEL, J.U. (1990). "Calidad en el centro escolar", en *Cuadernos de Pedagogía*, nº 186, pp. 75.

EUROPEAN FOUNDATION FOR QUALITY MANGEMENT (1994). *The european Quality Award: Aplication Brochure.* Tilburg (The Netherlands): Pabo Print.

EUROPEAN FOUNDATION FOR QUALITY MANAGEMENT (1995). *Autoevaluación: directrices.* Madrid: Club de Gestión de Calidad.

EURYDICE (1997). *Una década de reformas en la educación obligatoria de la Unión Europea* (1984-1994). Bruselas: CE.

FARLEY, J. y otros (1985). *Reconceptualization of vocational education program evaluation.* Eniversity Ohio.

FERNÁNDEZ, B. (1990). *Modelo causal de evaluación de centros.* Tesis doctoral. Universidad de Oviedo.

FERNÁNDEZ, J. (1990). "Evaluación de la calidad docente de la Universidad Complutense (1989-1990)" en *Gaceta Complutense*, Madrid, (1990), n. 74; p. 12-15.

FERNÁNDEZ-BALLESTEROS, R. (1995). *Evaluación de programas. Una guía práctica en ámbitos sociales, educativos y de salud.* Madrid: Edit. Síntesis.

FINDLAY, P. (1990). "Developments in the Perfomance Indicator Debate in the United Kingdom". En GOEDEGEBUURE, L.C.J., MAASSEN, P.A.M. y

WESTERHEIJDEN, D.F. (Eds). *Peer Review and Perfomance Indicators*, Utrecht, NL: Uitjeverij. Lemma.

FINNEY, J. and MOOS, R.H. (1989). "Teoría y método en evaluación de tratamientos", en *Evaluation and Program Planning*. V. 12, nº 14, pp. 367-375.

FIRESTONE, W. y HERRIOT, R.E. (1982). "Prescripcions for effective elementary schools do not fit secuntary schools", en *Educational Administation Quaterly*, 18, (2), págs. 39-59.

FUENTES, A. (1988). *Procesos funcionales y eficacia de la escuela. Un modelo causal*. Madrid: Universidad Complutense.

FUEYO GUTIÉRREZ, A. y FERNÁNDEZ RAIGOSO, M. (1989). "Análisis de la documentación sobre evaluación de centros en España", en *Revista Investigación Educativa*. Barcelona, n. 13; p. 183-198.

FULLAN, M. (1985). "The management of change", en HOYLEE, E. y MCMAHON, A. *The management of schools*. London: Kogan Pape.

GARCÍA, J.M. (1995). "La evaluación institucional a través del departamento de recursos humanos en una institución universitaria privada", en *Bordón, 47*, 1, 17-30.

GARCÍA CARRASCO, J. (1988). "El problema de la evaluación del profesorado universitario", en *Studia Paedagogica*, Salamanca, n. 20; p. 27-34.

GARCÍA CUETO, E. (1993). *Introducción a la psicometría*. Madrid: Siglo XXI.

GARCÍA GARRIDO, J.L. (1981). "La calidad de la educación mundial". Escuela Asturiana de Estudios Hispánicos, *La Calidad de la Educación. Exigencias científicas y condicionamientos individuales y sociales*. Madrid: Consejo Superior de Investigaciones Científicas, pp. 25-35.

GARCÍA GARRIDO, J.L. (1992). "El profesorado universitario: un análisis comparativo de los modelos de formación". *Actas del Congreso Internacional de Universidades*.

GARCÍA GARRIDO, J.L. (1996). *Diccionario europeo de la educación*. Madrid: Dykinson.

GARCÍA HOZ, V. (1975). "Una pauta para la evaluación de centros educativos", en *Revista Española de Pedagogía*, número 130.

GARCÍA MAZA, D. (1997). "Del balance social, al balance ético", en *La rentabilidad de la ética en la empresa*. Madrid: Visor-Argentaria.

GARCÍA RAMOS, J.M. (1992). "Evaluación de la eficacia docente", en *Bordón*, 43 (4).

GARCÍA RAMOS, J.M. (1992). "Recursos metodológicos en la evaluación de programas", en *Bordón*, 43(4), pp. 461-476

GARCÍA RAMOS, J.M. (1995). "La evaluación institucional a través del Departamento de Recursos Humanos en una institución universitaria privada", en *Bordón*, 47 (1).

GARCÍA RAMOS, J.M. (1997). "Análisis factorial confirmatorio en la validación del constructo competencia docente del profesor universitario", en *Bordón*, 49 (4).

GARCÍA RAMOS, J.M. (1997). "Valoración de la competencia docente del profesor universitario. Una aproximación empírica", en *Revista Complutense de Educación*, 8 (2). Madrid.

GARCÍA RAMOS, J.M. y CONGOSTO LUNA, E. (1996). "Un sistema de evaluación institucional como elemento de promoción de la calidad educativa en un centro universitario privado", en *Studia Pedagogica*. Salamanca, pp. 123-158.

GENTO PALACIOS, S. (1996). *Instituciones educativas para la calidad total*. Madrid: La Muralla.

GEORGE, M.D. (1982). "Assessing Program Quality". En WILSON, R. (Ed.) *Designing Academic Program Review*. San Francisco: Jossey Bass.

GIBBS, G. (1995). "Quality in Research and Quality in Teaching". En *Actas de las III Jornadas del Programa de Formación para la función docente del profesorado universitario de la Universidad Politécnica de Cataluña: La Calidad en la enseñanza universitaria*. Barcelona. Mayo 1995, pp. 7-21.

GIL, F. (1988). "Análisis de los conceptos de coherencia docente y coherencia del centro como indicadores de la calidad educativa". En las *Actas del IX Congreso Nacional de Pedagogía*, pp. 270.

GIL TRAVER, F. (1999). *Indicadores sobre entorno y procesos. Perspectiva española en el marco del proyecto de Indicadores de los sistemas educativos de la OCDE*. Madrid: INCE.

GIMENO SACRISTÁN, J. y PÉREZ GÓMEZ, A. (1989). *La enseñanza: su teoría y su práctica*. Madrid: Akal.

GINSBURG, A. (1989). "Revitalizing Program Evaluation", en *Evaluation Review*. V. 13, nº 6, pp. 359-398.

GLASER, R. (1963). "Instructional Technology and the measurement of learning outcoes", en *American Psychology*, 18: 519-521.

GLASER, R. y SILVER, E. (1994). "Assessment, Testing and Instructions: Retrospect and Prospect", en *Review of Research in Education*, 20: 393-423.

GOBANTES, J.M. (2000). *Los Centros del Profesorado: Evaluación y Cambio Educativo*. La Laguna: Universidad de La Laguna. Servicio de Publicaciones. ISBN: 84-699-1661-0.

GOETZ, J.P. y LECOMPTE, M.D. (1988). *Etnografía y diseño cualitativo en investigación educativa*. Madrid: Morata.

GOMEZ DACAL, G. (1979). "Evaluación de la calidad de la enseñanza", en *Vida Escolar*, 170-1, 40-45.

GONZÁLEZ, M.J. (1999). *La universidad del siglo XXI. Libertad, competencia y calidad*. Madrid: Círculo de empresarios.

GONZÁLEZ RAMÍREZ, T. (2000). "Modelos de Análisis secuencial en inspección educativa". En MATAS TERRÓN, A.; GARCÍA JIMÉNEZ, E.; GIL FLORES, T. y GONZÁLEZ RAMÍREZ, T., *Análisis de datos II*. Sevilla: Kronos.

GONZÁLEZ SOLER, A. (1987). "Evaluación de centros y programas educativos", en *Revista Comunidad Educativa*, nº 151, Madrid, ICCE, pp. 6-9.

GONZÁLEZ TIRADO, R.Mª. (1982). "Un modelo de análisis y diagnóstico de la calidad de enseñanza", en *Apuntes de Educación*, 5, 5-6.

GOODMAN, O. y PENNINGS, J. (1997). *News perspectives on organizational effectiveness*. San Francisco: Jossey-Bass.

GUBA, E.G. y LINCOLN, Y.S. (1982). *Effective evaluation*. California. Jossey-Bass publishers.

GUBA, E. y LINCOLN, Y. (1989). *Fourth generation evaluation*. California: Sage. Newbury Park.

GUERCI, B. (1993). *La cuestión de los fines en educación*. San Salvador de Jujuy: Secretaría de ciencia y técnica y estudios regionales.

GUILFORD, J.P. (1936). *Psychometric Methods*. New York: McGraw-Hill.

HAERTEL, E. (1991). "New Forms of Teacher Assessment", en *Review of Research in Education*, 17: 3-29.

HALL, G. (1987). "The principal's role in setting school climate", en *American Educational Research Association*. Washington.

HALPIN, A.W. y CROFT, D.B. (1963). *The organizational climate of Schools*. Chicago: Midwest Administration Center. University of Chicago.

HAMBLETON, R. (1994). "The rise and fall of criterion-referenced measurement", en *Educational Measurement: Issues and Practice*, 13, 4: 21-26.

HAMBLENTON, R. y SIRECI, S. (1997). "Future directions for norm-referenced and criterion-referenced achievement testing", en *International Journal of Educational Research*, 27, 5: 379-393.

HERNÁNDEZ PINA, F. (1995). *Bases metodológicas de la investigación educativa. Fundamentos*. Murcia: D.M.

HERNÁNDEZ PINA, F. (1995a). "La evaluación de centros educativos". *Actas del Congreso Nacional sobre Orientación y Evaluación Educativa*. A Coruña.

HERNÁNDEZ PINA, F. (1995b). "La evaluación de los alumnos en el contexto de la evaluación de la calidad de las Universidades". *Revista de Investigación Educativa*, 2, (14), 25-50.

HOLDAWAY, E.A. (1988). "Institutional Self-Study and the improvement of quality". En KELLS, H.R. Y VAN VUGHT, F. (Eds). *Self regulation, self-study and Program review in Higher Education*. Culemborg. NL: Lemma.

HOLLY, P. y HOPKINS, D. (1988). "Evaluation and School Improvement", en *Cambridge Journal of Education*, Vol. 18, nº 2. pp. 221-245.

HOUSE, E.R. (1988). "Tres perspectivas de la innovación educativa: Tecnológica, Política y Cultural", en *Revista de Educación*, nº 1, 286, pp. 5-34.

HUFNER, K. y RAU, E. (1987). "Measuring perfomance in higher education: problems and perspectives", en *Higher education in Europe*, 12, 4, 5-13.

IBÁNEZ-MARTÍN, J.A. (1990). "Dimensiones de la competencia profesional del profesor de universidad", en *Revista Española de Pedagogía*, Madrid, n. 186; p. 239-257.

ICE (1995). *Los indicadores de rendimiento en la evaluación institucional universitaria*. Zaragoza.

INCE (1995). *Evaluación de la Educación Primaria*. Madrid: MEC.

INCE (1997). *Evaluación comparada de la enseñanza y el aprendizaje de la lengua inglesa: España, Francia, Suecia*. Madrid: MEC.

INCE (1997). *Evaluación de la educación primaria*. Madrid: MEC.

INCE (1998). *Diagnóstico del sistema educativo. La escuela secundaria obligatoria*. 1997. 6 tomos. Madrid: MEC.

ISAACS, D. (1977). *Cómo evaluar los centros educativos*. Pamplona: EUNSA.

JACOBSON, H.K. (1978). "Framework for evaluation: Indicators of effort, performance, effets". En JACOBSON, H.K. (Eds). *Evaluating Advancement Programs*. New Directions for institutional Advancement 1. San Francisco: Jossey-Bass.

JENNINGS, J.F. (1998). *Why National Standards and Tests? Politicis and the quest for better schools*. London: Sage.

JOFRE, L. y VILALTA, J. (1998). "El Plan Nacional de Evaluación de la Calidad de las Universidades. Los objetivos del Plan: Evaluación para mejorar", en *Revista de Educación, 315*, 97-108.

KELLS, H.R. (1988). *Self-Study Processes. A Guide for Post secondary and similar Service-Oriented Institutions and Programs*. 3ª edic. New York: American Council of Education.

KELLS, H.R. (1988a). "Perspectives on university self-assessment for western Europe". En KELLS, H.R. y VAN VUGHT, F.A. (Eds). *Self-regulation, selfstudy and program review in Higher Education*. Cullemborg, NL: Lemma.

KELLS, H.R. y VAN GUTH, F.A. (1988). "Theoretical and practical Aspects of a Self-regulation and Quality Control System for Duth Higher Education", en *Tijdschrift voor Hoger Onderwings*, 6, 1, 15-20.

KORETZ, D. (1992). "State and national assessment", enn M.C. ALKIN (Ed.) *Encyclopedia of Educational Research*. Nueva York: McMillan, vol. 4, 1262-1267.

LANDSHEERE, G. (1985). *Diccionario de evaluación y de la investigación educativa*. Barcelona: Oikos-Tau.

LANDSHEERE, G. DE (1996). *El pilotaje de los sistemas educativos*. Madrid: La Muralla.

LÁZARO MARTÍNEZ, A.J. (1922). "La formación de indicadores en evaluación", en *Bordón*, vol. 43, nº 4, pp. 477-495.

LE BOTERF, G. et al. (1993). *Cómo gestionar la calidad de la formación*. Barcelona: Ediciones Gestión 2000.

LÉVY-LEBLOND, J.M. (1991). "La physique, une science sans complexe?", en *Les théories de la complexité.* Fogelman,F., Seuil, París.

LEWIS, C. y SHEEBAN, K. (1990). "Using bayesian decision theory to design a computeired mastery test", en *Applied Psychological Measurement*, 14, 4: 367-386.

LINN, R. (1988). "Medición educativa: algunos problemas y tendencias actuales de la investigación educativa", en DENDALUCE, I. (Coord.). *Aspectos metodológicos de la investigación educativa*. Madrid: Narcea.

LINN, R. (Ed.) (1993). *Educational Measurement.* New York: Macmillan Pub. Co.

LOHMAN, D. (1997). "Lessons from the history of inteligence testing", en *International Journal of Educational Research*, 27, 5: 359-377.

LÓPEZ, J.A. y MORENO, M.L. (1997). *Resultados de Ciencias. Tercer Estudio Internacional de Matemáticas y Ciencias (TIMSS).* Madrid: MEC.

LÓPEZ FEAL, R. (1986). *Construcción de instrumentos de medida en Ciencias conductuales y sociales*. Barcelona: Alamex.

LÓPEZ MOJARRO, M. (1999). *A la calidad por la evaluación*. Madrid: Escuela Española.

LÓPEZ RUPÉREZ, F. (1994). *La gestión de calidad en la educación*. Madrid: La Muralla.

LÓPEZ RUPÉREZ, F. (1997). "Complejidad y Educación", en *Revista Española de Pedagogía*, 206, pp. 103-112.

LORD, F. (1980). *Aplications of item response theory to practical testing problems.* Hillsdade: Lawrence Erlbawm.

LORD, F. y NOVICK, M. (1968). *Statistical theories of mental test scores.* Reading M.A.: Addison-Wesley.

LOUCK-HORSLEY, S. et al. (1987). *Continuing to Learn. A Guidebook for Teacher Development.* The Regional Laboratory for Educational Improvement of the Northest and Island & The National Staff Development Council. Andover, Mass.

LUJÁN CASTRO, J. y PUENTE AZCUTIA, J. (1996). *Evaluación de los centros docentes El Plan EVA*. Madrid: MEC.

MACDONALD, B. (1983). "La evaluación y el control de la educación". SACRISTÁN y PÉREZ GÓMEZ: *La enseñanza: su teoría y su práctica*. Madrid: Akal, págs.: 467-478.

MACPHERSON, R.J.S. (1998). *The politics of accountability. Educative and international perspectives*. London: Corwin Press.

MADAUS, G.; ARASIAN, P. y KELLAGHAN, TH. (1980). *School Effectiveness: A reassessment of the evidence*. New York: McGraw-Hill.

MARCELO GARCÍA, C. (1989). *Introducción a la formación del profesorado. Teoría y métodos.* Sevilla: Universidad de Sevilla.

MARCHESI, A. y MARTÍN, E. (1998). *Calidad de la enseñanza en tiempos de cambio*. Madrid: Alianza editorial.

MARTÍN, M.O. (1996). *Science Achievement in the Primary school years: IEA's Third International mathematics and Science Study*. Chesnut Hill: Boston College.

MARTÍN RODRÍGUEZ, E. (1988). "Profesión docente y autoevaluación institucional", en *Revista de Educación*, Madrid, n. 285; p. 33-43.

MARTÍNEZ, F. (2000). *La institucionalización de la evaluación.* (Documento difundido por Internet). México. Inédito.

MARTÍNEZ ARIAS, M.R. (1995). *Psicometría: Teoría de los Test psicológicos y educativos*. Madrid: Síntesis.

MATEO, J. (1988). "Medición educativa: estado de la cuestión en el ámbito español". En DENDALUCE, I. (Coord.). *Aspectos metodológicos de la investigación educativa*. Madrid: Narcea.

MATEO ANDRÉS, J. (1995). "Análisis de instrumentos de evaluación de centros", en *Revista Galega de Psicopedagoxía*. Monográfico, nº. 6, 69-98.

MATEO ANDRÉS, J. (2000). "La evaluación del profesorado y la gestión de la calidad de la educación. Hacía un modelo comprensivo de evaluación sistemática de la docencia". *Revista de Investigación Educativa*, Vol. 18, nº 1, pp. 7-34.

MAYOR RUIZ, C. y GONZÁLEZ RAMÍREZ, T. (2000). *La mejora de los planes de estudio en la Universidad. Modelo de seguimiento para la calidad*. Vicerrectorado de Calidad. Universidad de Sevilla.

MCDONALD, B. (1971). "The evaluation of the humanities Curriculum Proyect: a holistic approach", en *Theory into practice*, 10 (3), págs. 163-167.

MCDONALD, B. (1983). "The evaluation of the humanities curriculum Proyect: a holistic approach", en *Theory into practice*, 3, 157-189.

MEC (1998). *Modelo Europeo de Gestión de la Calidad*. Madrid: MEC.

MEDINA RIVILLA, A. (1988). *Didáctica e interacción en el aula*. Madrid: Cincel.

MEDINA RIVILLA, A. y VILLAR ANGULO (Coord.) (1995). *Evaluación de programas, centros y profesores*. Madrid: Universitas.

MESSIK, S. (1994). "The Interplay of Evidence and Consequences in the Validation of Performance Assessment", en *Educational Research*, 23, 2: 13-22.

METFESEEL, N.S. y MICHEL, W.B. (1967). "A paradigm involviong multiple criterion mesaures for the evaluation of the effectiveness of school programs". *Educational and Psychological Measurement*, 27, págs. 931-943.

MICHAVILA, F. (1996). "Calidad y evaluación de las instituciones Universitarias". En ALLEN, J. y MORALES, G. (Eds). *La universidad del siglo XXI y su impacto social*. Universidad de las Palmas de Gran Canaria.

MICHAVILA, F. y CALVO, B. (1998). *La universidad española hoy. Propuestas para una política universitaria*. Madrid: Síntesis.

MIDKIFF, R.M. & BURKE, J.P. (1987). "An Action Research Strategy for Selecting and Conducting Program Evaluations", en *Psychology in the Schools*. Volume 24, April.

MILLER, R.I. (1990). *Major american higher education issues and challenges in the 1990s*. Londres: Jessica Kingsley Pub.

MINISTERIO DE EDUCACIÓN Y CIENCIA. (1983). *Ley de Reforma Universitaria*. Madrid: Ministerio de Educación y Ciencia.

MINISTERIO DE EDUCACIÓN Y CULTURA (1997). *Modelo Europeo de Gestión de Calidad*. Madrid: Secretaría General de Educación y Formación Profesional. Dirección General de Centros Educativos.

MOERKERKE, G. (1996). *Assessment for flexible learning*. Utrecht: Lemma b.v.

MOOS, R.M. y TRICKETT, E. (1974). "Classroom environment scale manual", en *Consulting Psychologist Press*. Palo Alto.

MORA, J.G. (1998). "La evaluación institucional de la universidad", en *Revista de Educación, 315*, 29-44.

MORA RUIZ, J.G. (1991). *Calidad y rendimiento en las instituciones universitarias*. Madrid: Consejo de Universidades.

MORIN, E. (1991). "De la complexité: Complexus", en *Les théories de la complexité*. pp. 283-296. Fogelman, F., Seuil, París.

MORTIMER, P. (1986). "Autoevaluación escolar", en GALTON, M.: *Cambiar la escuela, cambiar el curriculum*. Barcelona: Martínez Roca.

MOSS, P. (1992). "Shifting Conceptions of Validity in Educational Measurement: Implications for Performance Assessment", en *Review of Educational Research*, 62, 3: 229-258.

MOSS, P. (1996). "Enlarging the dialogue in Educational Measurement: voices from Interpretative Research Traditions", en *Educational Researcher*, 25, 1: 20-29.

MULLIS, I.V.S. (1996). *Mathematics Achievement in the Primary school years: IEA's Third International mathematics and Science Study*. Chesnut Hill: Boston College.

MUÑIZ, J. (1990). *Teoría de respuesta a los Items*. Madrid: Pirámide.

MUÑIZ, J. (1994). *Teoría Clásica de los Tests*. Madrid: Pirámide.

MUÑIZ, J. y HAMBLETON, R. (1992). "Medio siglo de Teoría de Respuesta al Ítem", en *Anuario de Psicología*, 52: 41-66.

MUÑOZ CANTERO, J.M. (1995). "Hacia una concreción del término calidad". *Congreso Nacional sobre Orientación y Evaluación educativa. Libro de actas.*

MUÑOZ CANTERO, J,.M.; ABALDE PAZ, E.; DE SALVADOR GONZÁLEZ, X y RODICIO GARCÍA, L. (1995). "Sistema de indicadores para la evaluación de centros", en *Revista Galega de Psicopedagoxía*. Monográfico nº 6, 217-250.

MUÑOZ CANTERO, J.M. et al. (1995). "Análisis de la evaluación de la docencia universitaria por los alumnos de la Universidad de La Coruña". En AIDIPE (Comp.). *Estudios de Investigación educativa e Intervención psicopedagógica*. Valencia: AIDIPE, 289-293.

MUÑOZ CANTERO, J.M et al. (Coord.) (1995). "Calidad y Evaluación de centros", en *Revista Galega de Psicopedagoxía*. Monográfico nº 6.

MUÑOZ CANTERO, J.M. et al. (1995). "El profesorado de EGB perfil, eficacia y calidad docente", en *Revista Galega de Psicopedgoxía*, 10-11 (7) 102-122.

MUÑOZ CANTERO, J.M. et al. (1995). "Una aproximación al estudio de la docencia de la calidad. El buen docente universitario", en *Studia Pedagógica*.

MUÑOZ CANTERO, J.M. y RODRÍGUEZ SUÁREZ, Mª. (1996). "La práctica evaluadora como un indicador de la calidad universitaria", en *Revista Galega de Psicopedagoxía*, nº 13 (9), 131-161.

MUÑOZ-REPISO, M. (1995). *Calidad de la educación y eficacia de la escuela. Estudio sobre la gestión de los recursos educativos*. CIDE: Madrid.

MUÑOZ-REPISO, M. (1996). "La calidad como meta", en *Cuadernos de pedagogía*, 246, 52-57.

NAVAS, M.J. (1994). "Teoría Clásica de los Tests versus Teoría de Respuesta al Ítem", en *Psicológica*, 15: 175-208.

NEVO, D. (1986). "The conceptualization of educational Evaluation". En E.R. HOUSE (Edit.). *New Directions in Educational Evaluation*. London: The Falmer Press, 15-30.

NEVO, D. (1997). *Evaluación basada en el centro. Un diálogo para la mejora*. Bilbao: Mensajero.

OAKES, J. (1986). *Education indicators. A guide for policy makers*. New Brunswick, N.J. *Center for Policy Research in Education*.

OCDE (1990). *Escuelas y calidad de la enseñanza. Informe internacional*. Madrid: MEC-Paidós.

OCDE (1991). *Escuelas y calidad de enseñanza. Informe internacional*. Madrid: Paidós.

OCDE (1994). *La Qualité de l'enseignement*. París: CERI.

OCDE (1998). *Education at a glance. OCDE Indicators 1998*. Madrid: Mundiprensa libros.

OCDE (1999). *Education Policy Analysis. 1999*. Madrid: Mundiprensa libros.

ORDEN, A. DE LA (1988). "Calidad de los Centros Educativos, asunto para un Congreso", en *Bordón*, 40, 2.

ORDEN HOZ, A. DE LA (1985). "Hacia una conceptualización del producto educativo", en *Revista Investigación Educativa* nº 6.

ORDEN HOZ, A. DE LA (1985). "Investigación evaluativa". En ORDEN, A. (Ed.). *Investigación Educativa. Diccionario Anaya de Ciencias de la Educación*. Madrid: Anaya.

OROVAL, E. (Ed.). *Planificación, evaluación y financiación de sistemas educativos*. Madrid: Civitas.

OWENS, T.R. (1973). "Educational evaluation by adversary proceedings". En HOUSE, E.R. (Comp.) *School evaluation: the politics and process*. Berkeley: McCutchan.

PARLETT, M. y HAMILTON, D. (1977). "Evaluation ilumitaion". En GLASS, G.Y. *Evaluation studies review annual*, 1, págs. 140-157. Beverly Hills (Cal.): Sage.

PARLETT, M. y HAMILTON, D. (1983). "La evaluación como iluminación". En SACRISTÁN y PÉREZ GÓMEZ: *La enseñanza: su teoría y su práctica*. Madrid: Akal. págs.: 450-466

PASCUAL, P. (1991). "Evaluation and the improvement of quality", *en CRE-Action* 96, 95-102.

PAWSON, R. y TILLEY, N. (1997). *Realistic Evaluation*. London: Sage Publications.

PÉREZ, F. (1998). "Causas y consecuencias de la evaluación de las universidades: Para qué debe servir", en *Revista de Educación, 315*, 109-124.

PÉREZ JUSTE, R. (1995). "Evaluación de centros educativos" Enfoques y opciones metodológicas", en *Revista Galega de Psicopedagoxía*. Monográfico nº. 6, 7-43.

PÉREZ JUSTE, R. (1995). "Evaluación de programas educativos". En MEDINA RIVILLA, A. y VILLAR ANGULO, L.M. *Evaluación de programas educativos, centros y profesores*. Madrid: Universitas.

PÉREZ JUSTE, R. (1998). "Calidad de los Centros Educativos". *Documentos del curso a distancia sobre Modelo Europeo de Calidad Total*. Ite-CECE-UNED.

PÉREZ JUSTE, R. y GARCÍA RAMOS, J.M. (1989). *Diagnóstico, evaluación y toma de decisiones*. Madrid: Rialp.

PÉREZ JUSTE, R. y MARTÍNEZ DE ARAGÓN, L. (1989). *Evaluación de centros y calidad educativa*. Madrid: Cincel.

PERKINS, J.A. (1966). *The University in Transition*. Princeton: Princeton University Press.

PETERS, T.J. y WATTERMAN, R.H. (1992). *En busca de la excelencia: lecciones de las empresas mejor gestionadas en Estados Unidos*. Barcelona: Folio.

POPHAN, W.J. (1980). *Problemas y técnicas de evaluación educativas*. Madrid: Anaya.

POPHAN, W.J. (1981). *Modern educational measurement*. Englewood Cliffs (NJ): Prentice Hall.

PURKEY, S. y SMITH, M. (1983). "Effective School: a review", en *The elementary School Journal*, 83, 427-452.

QUINTANILLA, M.A. (1996). "Nuevas ideas para la universidad". En ALLEN, J. y MORALES, G. (Eds). *La universidad del siglo XXI y su impacto social*. Universidad de las Palmas de Gran Canaria.

QUINTANILLA, M.A. (1998). "El reto de la calidad en las universidades". En PORTA, J. y LLADONOSA, M. (Coords.). *La universidad en el cambio de siglo*. Madrid: Alianza.

QUINTANILLA, M.A. (1998). "En pos de la calidad: Notas sobre una nueva frontera para el sistema universitario español", en *Revista de Educación, 315*, 85-95.

QUINTAS, G. (Coord.) (1996). *Reforma y evaluación de la universidad*. Valencia: Universitat de València.

RILEY, K.A. y NUTTALL, D.L. (Eds.) (1994). *Measuring Quality. Education Indicators*. London: The Falmer Press.

RODRÍGUEZ, S. (1997). "La evaluación institucional universitaria", en *Revista de Investigación Educativa, 15 (2)*, 179-214.

RODRÍGUEZ, S. (1998). "El proceso de evaluación institucional", en *Revista de Educación, 315*, 45-65.

RODRÍGUEZ, S. (1998). "La evaluación institucional en España: Análisis y planteamientos futuros". En DE LUXÁN, J.M. (Ed.). *Política y reforma universitaria*. Barcelona: CEDECS.

RODRÍGUEZ, G.; GIL, J. y GARCÍA, E. (1996). *Metodología de la investigación cualitativa* Málaga: Aljibe.

RODRÍGUEZ DÍEZ, B. (1993). "Modelo de evaluación externa de la eficacia de los centros educativos", en *Revista Investigación Educativa*, Barcelona, n. 21; p. 7-27.

ROGER, S.; CISERO, C. y CARLO, M. (1993). "Techniques and Procedures for Assessing Cognitive Skills", en *RER*, 63: 201-243.

ROJO, J.M. (1993). "La autonomía universitaria y la búsqueda de la calidad". En BURGEN, A. *Metas proyectos de la educación superior: una perspectiva internacional*. Madrid: Fundación Universidad-Empresa.

ROSALES, C. (1999). *Evaluar es reflexionar sobre la enseñanza*. Madrid: Narcea.

ROTGER, B. (1989). *Evaluación formativa*. Madrid: Cincel.

RUIZ, J.Mª. (1998). *Cómo hacer una evaluación de centros educativos*. Madrid: Narcea.

RUL, J. (1995). *La memoria evaluativa del centro educativo. Un modelo integral de evaluación organizativa y curricular*. Generalitat de Catalunya. Colecciones de Gestión.

RUTTER, M. et al. (1979). *Fifteen thousand hours: secondary schools and their effects on children*. Cambrige: Harvard University Press.

RUTTER, M. (1983). "School effects on pupil progress: research findings and polity implications". En SCHULMAN, L. y SYKES, G. (Eds.). *Handbook of teaching and policy*. New York: Longman.

SALMERÓN, H. (1989). "El C.A.E.S., un cuestionario piloto sobre evaluación de calidad de los espacios arquitectónicos escolares". En las *Actas del IX Congreso Nacional de Pedagogía*, pp. 259-260.

SALVADOR, L. (1997). "Evaluación institucional universitaria en Europa y América Latina", en *Revista Española de Pedagogía, 208*, 429-444.

SANTOS GUERRA, M.A. (1990). *Hacer visible lo cotidiano*. Madrid: Akal.

SANTOS GUERRA, M.A. (1991). "Estrategias para la evaluación interna de los centros educativos", en *Curso de Formación para Equipos Directivos: Cuaderno* nº 9. Madrid: Ministerio de Educación y Ciencia.

SCHEERENS, J. (1990). "School effectiveness and the development of process indicators of school funtioning". En SCHEERENS, J. *School Efectiveness and school Improvement*. Lisse (Holanda): Swets&Zeitlinger, pág. 61-80.

SCHEERENS, J. (1991). "Process indicators of school functioning: a selection based on the research literature on school effectiveness", en *Studies in Educational Evaluation*, 17, 371-403.

SCHEERENS, J. (1992). *Effective Schooling*. London: Cassell.

SCHMIDT, W.; JAKWERTH, P. y MCKNIGHT, C. (1998). "Curriculum sensitive assessment: content does make a difference", en *International Journal of Educational Research,* 29: 503-527.

SCHOEMOKER, J. y FRASER, H. (1987). "What principals can do: some implications from studies of effective schooling", en *Phi delta Kappan*, 63.

SCRIVEN, M. (1967). "The Methodology of Evaluation". En STAKE, R.E. (Comp.). *Curriculum Evaluation. AERA Monograph Series on Curriculum Evaluation (v.1)* Chicago: Rand McNally.

SEGERS, M. (1996). "Assessment in a Problem-based Economics Curriculum". En BIRENBAUN, M. y DOCHY, F. (Eds.). *Alternatives in Assessment of Achievements Learning Process and prior Knowledge*. Boston: Kluwer Academic Publishers.

SELDEN, R.W. (1990). "Developing educational indicators: a state national perspective", en *International Journal of Education Research*, 66, 383-393.

SERRANO, J. y otros (1995). "Aportaciones recientes en medición y evaluación educativas", en *Revista de Investigación Educativa*, 26: 129-158.

SHAVELSON, R.J.; McDONNELL, L.M. y OAKES, J. (1991). "What are educational indicators and indicator systems?", en *American Institutes for Research. ERIC-EDO-TM-91-5.*

SIMONS, H. (1985). "Againts the rules: procedure problems in self-evaluation", en *Curriculum Perspectives*. Vol.5, nº 2 Octubre, p.1-6.

SIMONS, H. (1995). "La autoevaluación escolar como proceso de desarrollo del profesorado: En apoyo a las escuelas democráticas", en *Volver a pensar la educación. Prácticas y discursos educativos. Vol. II.* Madrid: Morata.

SIZER, J. (1979). "Assessing institutional perfomance and overviw", en *International Journal of Institutional Management in Higher Education, 3, 1*, 49-77.

SIZER, J. (1982). "Assessing institutional Perfomance and Progress". En WAGNER, L. (Eds). *Agenda for instituional change in Higher Education*. Guilford, UK: Society for Research in Higher Education.

SLEDSENS, G. et al. (1996). "L'evaluations des systemes educatifs", en *Scientia Paedagogica Experimentalis*, 33, 181-202.

STAKE, R. (1987). "Evaluation Design, Instrumentation, Data Collection, and Analysis of Data". En WORTHEN & SANDERS: *Educational Evaluation: Theory and Practice*. Charles A. Jones Publication. Wadsworth Publishing Company, Inc.

STAKE, R.E. (1975). *Program evaluation, particulary responsive evaluation*. Teachers College Record, 68, págs. 523-540.

STAKE, R.E. (1983). "La evaluación de programas, en especial la evaluación de réplica". En W.B. DOCKRELL y D. HAMILTON: *Nuevas reflexiones sobre la investigación educativa*. Madrid: Narcea.

STEVENS, S. (1946). "On the theory of scales of measurement", en *Science*, 103: 677-680.

STEVENS, S. (1951). "Mathematics, measurement and psychometrics", en STEVENS, S. (Eds.). *Handbook of Experimental Psychology*. New York: John Wiley.

STENHOUSE, L. (1984a). "Evaluating Curriculum Evaluation". En C. ADELMAN: *The Politics and Ethics of Evaluation*. London: Croom Helm.

STENHOUSE, L. (1984b). *Investigación y desarrollo del currículum*. Madrid: Morata.

STOEL, W. y SCHEERENS, J. (1988). *The Stability of school effects across contexts and nations*. AERA. Convetion, New Orleans.

STUFFLEBEAM, D. y SHINKFIELD, A.J. (1987). *Evaluación sistemática: guía teórica y práctica*. Madrid: Paidós/MEC.

SUCHMANNN, E.A. (1967). *Evaluative Research*. New York: Russell Sage Foundation.

TEJEDOR, F.J. (1997). "La evaluación institucional en el ámbito universitario", en *Revista Española de Pedagogía, 2,8* , 413-428.

TEJEDOR, J.; GARCÍA-VALCÁRCEL, A y RODRÍGUEZ CONDE, M.J. (1993). "Perspectivas metodológicas actuales de la evaluación de programas en el

ámbito educativo". *VI Seminario de modelos de investigación educativa*. Madrid. 23-25 septiembre de 1993.

TEJEDOR TEJEDOR, F.J. (1990). "Indicadores de la calidad docente para la evaluación del profesor universitario", en *Revista Española de Pedagogía*. Madrid (1990), n.186; p. 259-279.

TEJEDOR TEJEDOR, F.J. (1990). "La Evaluación del profesorado en la Universidad de Santiago", en *Revista Española de Pedagogía*, Madrid, n. 186; p. 337-361.

TEJEDOR TEJEDOR, J. (1999). "La evaluación institucional como estrategia para mejorar la calidad de la enseñanza". *Conferencia: Jornadas sobre evaluación*. La Coruña.

TIANA, A. (1993). "Evaluación de centros y evaluación del sistema educativo", en *Bordón*, Vol. 45, nº 3. pp: 283-293.

TIANA, A. (1998). "La evaluación externa de los centros educativos: posibilidades y limitaciones". *Ponencia V Jornadas Pedagógicas de la Federación Navarra de Ikastolas*. Pamplona.

TIANA, A. y MUÑOZ, F. (1997). "Indicadores educativos", en *Cuadernos de Pedagogía*, 256, 49-76.

TORANZOS, L.V. (1997). El problema de la calidad en el primer plano de la agenda educativa. *Actas I Reunión del Comité ejecutivo del programa de Evaluación de la Calidad de la Educación*. Buenos Aires: OEI/Ministerio de Cultura y Educación Rep. Argentina.

TORGERSON, W. (1958). *Theory and Methods of Scaling*. New York: Wiley.

TYLER, L.L. (1967). "Changing conceps of educational evaluation". En STAKE (Comp.). *Perpectives of curriculum evaluation*. New York: Rand McNally.

TYLER, R.W. (1977). *Principios básicos del currículo*. Buenos Aires: Troquel.

TYLER, R.W. (1991). "National Assessment of Educational Progress (NAEP)" En T. HUSEN y T.N. POSTLETHWAITE. *Enciclopedia Internacional de la Educación*. Barcelona: MEC-Vicens-Vives, pp. 4194-4196.

TYLOR, F.W. (1947). *Scientific Management*. New York. Harper.

UNESCO (1998). *Informe mundial sobre la educación*. Madrid: Santillana/ed. UNESCO.

UGALDE, M. (1995). "Aplicación de la mejora continua a la gestión de un programa de formación", en *Organización y Gestión Educativa*, 3, pp. 41-48.

UNIVERSIDAD DE CASTILLA LA MANCHA (1999). *II Foro de reflexión sobre el Plan Nacional de Evaluación de la Calidad de las Universidades*. Universidad Castilla La Mancha. Oficina de Evaluación de la Calidad.

UNIVERSIDAD NACIONAL DE CUYO (1996). *Evaluación institucional: autoevaluación y evaluación externa*. Mendoza: Universidad Nacional de Cuyo.

VALERO CRESPO, M. (1988). "La evaluación del profesor a través de los alumnos" en *Educa*, Cádiz, 4, n. 14; p. 26-30.

VAN DE BERGHE, W. (1997). *Aplicación de las Normas ISO 9000 a la enseñanza y la formación. Interpretación y orientaciones desde una perspectiva europea*. Luxemburgo: CEDEFOP.

VAN DER LINDEN, J. y EGGEN, T.J. (1986). "An empirical bayerian approach to item banking", en *Applied Psychological Measurement*, 10: 345-354.

VAN DER MOLER, H.J. (1999). "Creación, trasmisión y aplicación del conocimiento a través del sistema de educación superior". En BURGEN, A. *Metas y proyectos de la educación superior: una perspectiva internacional*. Madrid: Fundación Universidad-Empresa.

VAN DE VEN, A. y FERRY, D. (1980). *Measuringand assesing organizations*. New York: Wiley.

VAZQUEZ, G. (1981). "Límites de la investigación sobre la calidad de la educación". Escuela Asturiana de Estudios Hispánicos, *La Calidad en la Educación. Exigencias científicas y condicionamientos individuales y sociales*. Madrid: Consejo Superior de Investigaciones Científicas, pp. 43-57.

VÉLAZ DE MEDRANO URETA, C. et al. (1995): *Evaluación de programas y de centros educativos. Diez años de investigación*. Madrid. MEC-CIDE.

VIEIRA, J.T. (1997). "Evaluación institucional: Objetivos y criterios", en *Revista Española de Pedagogía, 208*, 445-458.

VILLAR, L.M. y MARCELO, C. (1991). "La evaluación de los climas socio-educativos en el ámbito universitario". En MEDINA RIVILLA, A. (Coord.). *Teoría y métodos de evaluación*. Madrid: Cincel.

VILLAR ANGULO, L.M. (1988). *Conocimiento, creencias y teorías de los profesores*. Alcoy: Marfil.

VILLARROEL, C. (1999). *Mitos y leyendas de la evaluación y acreditación universitarias*. (Documento difundido por Internet). Cumaná. Inédito.

VROEIJENSTI, T. y ACHERMAN, H. (1991). "Evaluación de la calidad basada en el control frente a la evaluación de la calidad basada en la mejora". En DE MIGUEL, M., MORA, J.G. y RODRÍGUEZ, S. *La evaluación de las instituciones universitarias*. Madrid: Consejo de Universidades. Secretaría General.

WAINER, H. (1990). *Computerized Adaptative Testing: A Primer*. Hillsdale, N.Y. Lawreces Erlbaum.

WEIS, J. (1989). "Evaluation as Subversive Educational Activity". En GOEFFREY MILBURN, IVOR F. GOODSON Y ROBERT J. CLARK (Eds.). *Re-Interpreting Curriculum Research: Images and Arguments*. Ontario: The Falmer Press.

WILLIAMS, J. y WATSON, L. (1995). "Gestión de la Calidad Total (TQM) en Educación", en *Organización y Gestión Educativa*, 3, pp. 8-13.

WILSON, J.D. (1992). *Cómo valorar la calidad de la enseñanza*. Madrid: Paidós. MEC.

WOLF, R.L. (1974). "The citizen as jurist: A new mode of educational evaluation", en *Citizen Action in education*, 4.

WOLF, R.L. (1975). "Trial by jury: a new evaluations method", en *Phi Delta Kappan*, 57, págs. 185-187.

WOLF, R.L. (1990). *Evaluation in education*. New York: Praeger Pb.

WOLF, R.M. (1998). "Validity issues in international assessment", en *International Journal of Educational Research*, 29: 491-501.

WORTEN, B.R. & SANDERS J.R. (1987). *Educational Evaluation: Theory and Practice*. Charles A. Jones Publication. Wadsworth Publishing Company, Inc.

YORKE, D. (1987). "Indicators or institutional achievement: some theoretical and empirical considerations", en *Higher Education*, vol. 16, 1.

YOUNG, K. et al. (1983). *Ungerstanding Acreditation*. San Francisco: Jossey-Bass Publishers.

ZABALZA. M. (1995). "Las diez dimensiones de la calidad". *Ponencia presentada en el II Congreso Internacional de Educación Infantil: Investigaciones y Experiencias*. Córdoba, Marzo de 1995.

ZABALZA BERAZA, M.A. (1990). "Evaluación orientada al perfeccionamiento", en *Revista Española de Pedagogía*, Madrid, n. 186, p. 295-317.

REFERENCIAS LEGISLATIVAS

En este anexo ofrecemos algunas referencias legislativas en relación a dos de los ámbitos evaluativos que han sido tratados en el texto: evaluación del profesorado y evaluación de centros educativos. Referencias concretas en relación a la evaluación del sistema educativo o la evaluación institucional universitaria, ya han sido tratadas en el capítulo correspondiente.

En las dos últimas décadas, periodo de mayor expansión de la evaluación, la legislación sobre evaluación ha cambiado de manera sustantiva; la evaluación de centros y del profesorado en sus distintos niveles educativos, ha sido con diferencia, los ámbitos evaluativos que de manera reiterativa, se han visto sometidos a mayores cambios legislativos; éste el motivo fundamental por el que nos hemos centrado en ellos únicamente.

Su inclusión en este texto se ha realizado para que sirva como material de consulta y acercamiento a algunos de los capítulos tratados en el texto, así como orientación y complemento para que el profesorado pueda llevar a cabo los procesos evaluativos dentro del marco prescriptivo que marca la legislación. En este apartado se incluyen también órdenes ministeriales que hacen referencia a las políticas de calidad que se han puesto en marcha de forma transversal para dotar de significado a los cambios legislativos propuestos.

Legislación Nacional relacionada con la Evaluación del Profesorado

Ley Orgánica de Reforma Universitaria [L.R.U.] (Ley 11/1983, de 25 de agosto. B.O.E. del 1 de septiembre).

Ley Orgánica de Ordenación General del Sistema Educativo [LOGSE] (Ley 1/1990, de 3 de octubre. B.O.E. del 4).

Orden Ministerial 25.820 de 12 de noviembre de 1992 (B.O.E. del 21), *sobre evaluación en Educación Infantil.*

Orden Ministerial 25.821de 12 de noviembre de 1992 (BOE del 21), *sobre evaluación en Educación Primaria.*

Orden Ministerial 25.707 de 12 de noviembre de 1992 (B.O.E. del 20), *sobre evaluación de la Educación Secundaria Obligatoria.*

Orden de 2 de diciembre de 1994 por la que se establece el procedimiento para la evaluación de la actividad investigadora en desarrollo del Real Decreto 1086/1989, de 28 de agosto, sobre retribuciones del profesorado universitario (B.O.E. 289/94 de 3).

Resolución de 6 de noviembre de 1996, de la Dirección General de Enseñanza Superior-Presidencia de la Comisión Nacional Evaluadora de la Actividad Investigadora, por la que se establecen los criterios específicos en cada uno de los campos de evaluación (B.O.E. 280/96 de 20).

Resolución de 26 de octubre de 1995, de la Dirección General de Investigación Científica y Técnica-Presidencia de la Comisión Nacional Evaluadora de la Actividad Investigadora, por la que se establecen los criterios específicos en cada uno de los campos de evaluación (B.O.E. 274/95 de 16 de noviembre).

Resolución de 13 de diciembre de 1993, de la Secretaría de Estado de Universidades e Investigación, por la que se establece el procedimiento para la evaluación de la actividad investigadora realizada por los miembros de las escalas de profesores de investigación, investigadores científicos y colaboradores científicos del

Resolución de 20 de junio de 1990, del Consejo de Universidades, por la que se establecen los criterios generales para la evaluación de la actividad docente del profesorado universitario (B.O.E. 156/90 de 30).

Resolución de 26 de septiembre de 1989, del Consejo de Universidades, por la que se establecen los criterios generales de evaluación del profesorado universitario para la evaluación global establecida en la Disposición Transitoria Tercera del Real decreto 1086/1989, de 28 de agosto (B.O.E. 239/89 de 5 de octubre).

Legislación Autonómica relacionada con la Evaluación del Profesorado

Comunidad Autónoma de Andalucía: Orden de 4 de diciembre de 1993, por la que se regula, con *carácter experimental, el programa de evaluación de centros y el programa de evaluación de proyectos curriculares de centro* y se establecen los criterios generales para su desarrollo (B.O.J.A. 139/93 de 24).

Comunidad Autónoma de Andalucía: Orden de 6 de septiembre de 1996, sobre *evaluación de centros docentes de niveles no universitarios,* sostenidos con

fondos públicos, en la Comunidad Autónoma de Andalucía (B.O.J.A. 115/96 de 5 de octubre).

Comunidad Autónoma de Canarias: Orden de 20 de diciembre de 1995, por la que se regula la *implantación y el desarrollo del programa de mejora de la calidad educativa* en la Comunidad Autónoma de Canarias (B.O.C. 8/96 de 17 de enero de 1996).

Comunidad Autónoma de Canarias: Orden de 7 de noviembre de 1996, por la que se establece el *proyecto de evaluación interna y desarrollo de centros,* los criterios generales para su desarrollo y las bases para la selección de centros que se acojan a dicho proyecto (B.O.C. 149/96 de 27).

Comunidad Autónoma de Cataluña: Orden de 20 de octubre de 1997, por la que se regula la *evaluación de los centros docentes sostenidos con fondos públicos* (D.O.G.C. 2511/97 de 5 de noviembre).

Legislación relacionada con la Evaluación de Centros Educativos

Ley Orgánica 8/1985, de 3 de julio, reguladora del Derecho a la Educación (BOE de14).

Ley Orgánica 1/199:0 de 3 de octubre de Ordenación General del Sistema Educativo (BOE de14).

Ley Orgánica 9/1995, de 20 de noviembre, de la participación, la evaluación y el Gobierno de los Centros Docentes [LOPEGCE] (B.O.E. del 21).

Resolución de 1 de septiembre de 1992 de la Dirección General de Coordinación y de la Alta Inspección, por la que se aprueba el Plan General de Actuación del Servicio de Inspección Técnica de Educación para el curso 1992/93 (BOMEC del 7 de septiembre).

Circular 10/1992 de la Dirección General de Coordinación y de la Alta Inspección por la que se emiten instrucciones a los Directores provinciales del Departamento sobre el Plan de seguimiento de los centros evaluados en el curso 1991/92, según lo previsto en el Plan Eva.

Resolución de 10 de noviembre de 1992 de la Dirección General de Coordinación y de la Alta Inspección, por la que se aprueba el Plan de evaluación de centros para el curso escolar 1992/93 y se establecen los criterios generales para su desarrollo.

Circular 12/1992 de la Dirección General de Coordinación y de la Alta Inspección por la que se dictan instrucciones para el desarrollo del Plan de Evaluación de Centros docentes durante el curso 1992/93.

Instrucciones de 5 de septiembre de 1996, de la Dirección General de Centros Educativos, por las que se regula el desarrollo y aplicación del Plan anual de mejora de los centros públicos dependientes del M.E.C. (BOMEC de7 de octubre).

Orden 4.578 de 21 de febrero de 1996 sobre la evaluación de los centros docentes sostenidos con fondos públicos (B.O.E. del 29).

Resolución de 2 de septiembre de 1997, de la Dirección General de Centros Educativos, por la que se dictan instrucciones para el desarrollo y aplicación del Plan Anual de Mejora en los centros públicos dependientes del M.E.C (BOE del 16).

Resolución de 17 de mayo de 1988, de la de la Dirección General de Centros Educativos, por la que se dictan instrucciones para la implantación, con carácter experimental, del Modelo Europeo de Gestión de Calidad en los centros docentes (BOE del 2 de junio).

Orden de 9 de junio de 1998 por la que se establece el plan anual de mejora en los centros docentes públicos dependientes del Ministerio de Educación y Cultura y se dictan instrucciones para su desarrollo y aplicación (B.O.E. 141/98 del 13).


80000004068781



ISBN 84-95212-83-8
9 788495 212832